

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
КРИВОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ФАКУЛЬТЕТ МЕХАНІЧНОЇ ІНЖЕНЕРІЇ ТА ТРАНСПОРТУ  
КАФЕДРА “АВТОМОБІЛЬНИЙ ТРАНСПОРТ”

ПОЯСНЮВАЛЬНА ЗАПИСКА  
ДО МАГІСТЕРСЬКОЇ РОБОТИ

“ УЗАГАЛЬНЕННЯ ДАНИХ РОБОТИ НАВАНТАЖУВАЛЬНО-  
ТРАНСПОРТНОГО КОМПЛЕКСУ КАР’ЄРУ ”

ВИКОНАВ: СЕМЬОНОВ МАКСИМ МИХАЙЛОВИЧ

КЕРІВНИК: МОНАСТИРСЬКИЙ ЮРІЙ АНАТОЛІЙОВИЧ

Допущений до захисту

Зав. кафедрою АТ

« \_\_\_\_\_ » листопада 2024 р.

Ю.А.Монастирський

Кривий Ріг – 2024

КРИВОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ФАКУЛЬТЕТ МЕХАНІЧНОЇ ІНЖЕНЕРІЇ ТА ТРАНСПОРТУ  
Кафедра “ Автомобільний транспорт ”  
Освітня програма – Автомобільний транспорт

Затверджую:

Зав. кафедрою АТ

Ю.А.Монастирський

« » вересня 2024 р.

ЗАВДАННЯ

НА МАГІСТЕРСЬКУ РОБОТУ

*Семьонову Максиму Михайловичу*

1. Тема роботи “УЗАГАЛЬНЕННЯ ДАНИХ РОБОТИ НАВАНТАЖУВАЛЬНО-ТРАНСПОРТНОГО КОМПЛЕКСУ КАР’ЄРУ”

затверджена наказом по університету від «16» вересня 2024 р. № 838с

2. Термін здачі закінченої роботи «30» листопада 2024 р.

3. Вихідні дані до роботи: *наукові дослідження та публікації за темою роботи.*

4. Зміст розрахунково-пояснювальної записки: *Аналіз стану питання, Методика дослідження, Результати дослідження, Висновки, Список використаних джерел*

5. Перелік графічного матеріалу *Графіки та діаграми по результатах досліджень.*

6. Дата видачі завдання - 2024 р.

КАЛЕНДАРНИЙ ПЛАН

Пор.№	Назва етапів випускної роботи	Термін виконання етапів проекту (роботи)	Примітка
1	<i>Аналіз стану питання</i>	<i>10.09.24 – 01.10.24</i>	
2	<i>Методичні засади</i>	<i>02.10.24 – 20.10.24</i>	
3	<i>Дослідження та аналіз результатів</i>	<i>21.10.24 – 21.11.24</i>	
4	<i>Формулювання висновків</i>	<i>22.11.24 – 24.11.24</i>	
5	<i>Оформлення презентаційної частини</i>	<i>25.11.24 – 30.11.24</i>	

Завдання видав

Керівник

МОНАСТИРСЬКИЙ Ю.А.

Завдання прийняв

СЕМЬОНОВ МАКСИМ МИХАЙЛОВИЧ

# Зміст

Зміст .....	3
Анотація .....	4
1. Введення.....	6
2. Аналіз стану питання за результатами попередніх досліджень.....	8
3. Методика досліджень .....	11
3.1. Вихідні дані.....	11
3.1.1. Набір даних продуктивності машин .....	12
3.1.2. Набір даних CAT MineStar Edge.....	13
3.2. Попередня обробка даних .....	14
3.3. Аналіз даних .....	16
3.4. Статистичний аналіз .....	17
3.5. Визначення достовірності результатів.....	18
4. Результати та їх аналіз .....	22
4.1. Аналіз даних .....	22
4.2. Статистичний аналіз .....	28
4.3. Статистична обробка результатів.....	36
5. Висновки .....	40
Список літератури.....	43

## Анотація

У цьому дослідженні визначені показники продуктивності навантажувачів і кар'єрних вантажівок під час гірничих робіт. Використані дані, люб'язно надані компанією ТОВ з П Цеппелін Україна, які були зібрані в режимі реального часу від навантажувачів і кар'єрів, що працюють у кількох кар'єрах, щоб розширити сферу дослідження та усунути упередження. Наша модель вказує на взаємозв'язки між декількома змінними та їхній вплив на продуктивність під час операції. Аналіз даних також було застосовано до інструментів залучення до землі (GET) для визначення основних графіків профілактичного обслуговування, щоб мінімізувати вплив простою основного обладнання на виробництво. Завдяки аналізу даних навантажувача було виявлено, що ефективний час циклу складає приблизно від 35 до 40 с, що забезпечує більшу корисну навантаження. Алгоритм класифікатора дерева рішень створив модель 87.99% точність оцінки продуктивності завантажувача на основі повного аналізу даних. Виходячи з розподілу виробничих змінних між кожним типом навантажувача, який працює в схожих робочих умовах, Caterpillar 992K і 990K були найпродуктивнішими машинами. Ефективність виробництва порівнювалася до та після періодів технічного обслуговування землерийних знарядь на ковшах навантажувачів. З використанням записів про технічне обслуговування та виробництво цих інструментів було зроблено висновок, що не було помітних змін у середньому виробництві та відсотковій зміні вартості виробництва до та після днів технічного обслуговування. Ефективність виробництва порівнювалася до та після періодів технічного обслуговування землерийних знарядь на ковшах навантажувачів. З використанням записів про технічне обслуговування та виробництво цих інструментів було зроблено висновок, що не було помітних змін у середньому виробництві та відсотковій зміні вартості виробництва до та після днів технічного обслуговування. Ефективність виробництва

порівнювалася до та після періодів технічного обслуговування землерийних знарядь на ковшах навантажувачів. З використанням записів про технічне обслуговування та виробництво цих інструментів було зроблено висновок, що не було помітних змін у середньому виробництві та відсотковій зміні вартості виробництва до та після днів технічного обслуговування.

Ключові слова: гірничодобувна промисловість ; оптимізація кар'єру ; навантажувачі ; аналітика даних ; машинне навчання

# 1. Введення

Розвиток нових технологій відкрив ширший погляд на машини та датчики завдяки розширенню збору великих даних. Інтернет речей (IoT) закріпився в гірничодобувній промисловості, дозволяючи з'єднувати обладнання та програмне забезпечення для отримання даних у режимі реального часу про численні робочі частини: аналіз небезпек, управління автопарком, моніторинг стану, системи сигналізації та навіть оптимізація процесів [ 1 , 2 , 3 ]. Машинобудівні компанії, які розробляють машини для видобутку корисних копалин, такі як Caterpillar, тепер оснащують нові покоління машин апаратним забезпеченням, яке дозволяє збирати великі дані [ 4 , 5 ] . Нове обладнання підключає датчики на всій машині до інтерфейсів користувача та зберігає записи всіх показників. Потім нове з'єднання датчик–інтерфейс користувача використовує структуровані дані для виведення аналізу показників продуктивності, ефективності та загального стану роботи машини [ 6 ] .

Хоча ці технологічні вдосконалення існують, вони не завжди повністю використовуються компаніями в гірничодобувній промисловості. Гірничодобувна промисловість є зрілою галуззю, яка звикла до того, як все робилося в минулому, тому що це дешевше, ніж впровадження нових технологій у галузі, що постійно розвивається [ 1 , 7 ] . Технології можуть принести користь гірничодобувній промисловості, забезпечуючи прогнозний аналіз, щоб вони могли бути проактивними і не реагувати на небезпеки, такі як аварії на шахтах або поломки машин, які затримують роботу. Проактивність може допомогти передбачити погані умови праці для покращення здоров'я працівників, передбачити вихід з ладу машин через поломку деталей або витік рідини, що спричинить пошкодження конструкції, або навіть запобігти смертям, таким як прогнозований зсув на шахті Бінгем у 2013 році [ 8 , 9 ] . Окрім прогнозного аналізу, нова технологія також може допомогти, визначаючи оптимальну продуктивність машин, забезпечуючи

операторам виробничі цілі, що дозволяє досягти повного потенціалу роботи, щоб не відставати від попиту.

Мета цього дослідницького дослідження полягає в тому, щоб дослідити переваги збору великих даних і полегшити ідентифікацію ключових показників ефективності за допомогою двох програмних інструментів Caterpillar, Cat Productivity (версія 2.0, Caterpillar Inc., Ірвінг, Техас, США) і Cat MineStar Edge. (версія 1.5.20220804p2, Caterpillar Inc., Ірвінг, Техас, США). Ці дві частини програмного забезпечення новіші в галузі та повільно впроваджуються. Аналіз виробничих даних, виведених з машин, буде корисним для визначення оптимальних виробничих показників для встановлення цілей операцій, а також візуалізації переваг технічного обслуговування. У цьому дослідженні зібрані дані будуть використані для визначення цих оптимальних показників ефективності з метою збільшення річного виробництва. Одна з перевірених гіпотез полягала в тому, щоб побачити, чи існує оптимальна точка при менших часах циклу, ідеально в діапазоні від 30 до 45 с, у якому середнє корисне навантаження матиме вищу продуктивність, таким чином встановлюючи виробничі цілі для операторів. Інша гіпотеза оцінювала, чи доречно технічне обслуговування ковшів навантажувача GET, оскільки наразі оператори визначають частоту заміни. Остання гіпотеза стверджує, що машинне навчання дасть прогностичну модель типу навантажувача та виробничих значень, пов'язаних із цим навантажувачем. Регресійні моделі можна використовувати для пошуку додаткових зв'язків змінних, які ще не виділені. Остання гіпотеза стверджує, що машинне навчання дасть прогностичну модель типу навантажувача та виробничих значень, пов'язаних із цим навантажувачем. Регресійні моделі можна використовувати для пошуку додаткових зв'язків змінних, які ще не виділені. Остання гіпотеза стверджує, що машинне навчання дасть прогностичну модель типу навантажувача та виробничих значень, пов'язаних із цим навантажувачем. Регресійні моделі можна використовувати для пошуку додаткових зв'язків змінних, які ще не виділені.

## 2. Аналіз стану питання за результатами попередніх досліджень

Попередні дослідження досліджували оптимізацію взаємодії парку без використання технології, яка значною мірою покладається на рівняння, які забезпечують число на основі змінних, зібраних на основі візуальних припущень, а також ручних вимірювань. Matsimbe (2020) провів аналіз даних про взаємодію лопати та вантажівки в кар'єрі в Малаві, щоб визначити, чи можуть вони оптимізувати розмір автопарку, використовуючи різну кількість вантажівок для однієї лопати [10] . Використовуючи секундомір для вимірювання часу циклу, а також численні рівняння для розміру корисного навантаження, які спираються на судження користувача про те, наскільки заповнений ковш лопати, вони змогли визначити збільшення розміру парку з лопатою поточного розміру, що призвело до часу черги. кар'єрних вантажних автомобілів збільшується на 6,40 хв [10]. Nday та ін. виконали аналогічну роботу в шахті в Демократичній Республіці Конго, з ручними розрахунками, які враховували умови транспортних доріг і обладнання, а також досвід оператора, щоб скоротити тривалість циклу приблизно на 8% [11] . Нарешті, Samatemba et al. нещодавно використовував рівняння для виявлення коефіцієнта використання, рівня виробництва, доступності обладнання, ефективності та продуктивності для кар'єрних вантажівок, навантажувачів і бурових установок у Південній шахті Чібулума в Замбії [12] . За допомогою простого аналізу ці дослідники змогли визначити, що всі їхні машини працювали з показником ефективності менше 50%, що призвело до значної втрати доходу [12] .]. Уявіть собі, якби їхні машини пов'язали датчики з програмним забезпеченням, яке могло б точно вимірювати корисне навантаження, тривалість циклів навантажувача та самоскида, а також час стояння вантажівок у черзі. Без використання рівнянь, які включають думку



користувачів, що створює ризик потенційної помилки користувача, вони могли б надати більш точне уявлення про те, як їх парк взаємодіє з лопатами та парками різних розмірів. Спільним для трьох шахт є те, що вони знаходяться в країнах третього світу, не маючи широкого доступу до технології підключення своїх машин до програмного забезпечення. У майбутньому доступ до технологій стане більш економічно ефективним і доступним для країн, що розвиваються, що призведе до потенційної еволюції впровадження з часом.

Включення датчиків і даних у машини дозволяє частіше використовувати стратегії прогнозованого технічного обслуговування. У минулому, без доступу до технології, щоб надати інформацію про показники продуктивності та стан машини, технічне обслуговування відбувалося, коли машини виходили з ладу або коли оператори використовували візуальні чи звукові сигнали, щоб визначити, що пошкодження вже відбулося [13]. Додавання датчиків і алгоритмів до машин може допомогти обійти це реактивне планування технічного обслуговування. Обладнання є більш надійним, що означає менше простоїв, коли деталі ламаються, і є помітне зниження витрат, оскільки менше одиниць обладнання ремонтується [14]. Такі алгоритми, як у машинному навчанні, можна використовувати для створення прогнозних моделей на основі минулих даних для прогнозування та планування технічного обслуговування машин. Басрі та ін. провели огляд цього і виявили, що завдяки комп'ютерному підходу до прогнозування несправностей машин компанії досягли кращої продуктивності та продуктивності з часом порівняно з реактивними методами обслуговування [15]. Через поточні збої в ланцюжку постачання для операцій життєво важливо передбачити, коли машини потребуватимуть запчастини, замість того, щоб чекати, поки машина вийде з ладу та потенційно втратить машину на кілька місяців, чекаючи на запасні частини.

Машинне навчання є популярною темою в багатьох галузях досліджень, а останні розробки в технологіях дозволяють легко отримати

доступ і прийняти. Дослідники в гірничодобувній промисловості зосередилися на цій тенденції та проводять постійно зростаючі дослідження, щоб побачити можливості цих алгоритмів разом зі збором великих даних. Нобахар та ін. використовували п'ять алгоритмів: лінійна регресія, дерево рішень, K-найближчі сусіди, випадковий ліс і посилення градієнта для моделювання операцій для оптимізації вибору парку [ 16] .]. Дослідження показало, що регресорний алгоритм градієнтного підсилення міг точно передбачити найкращий вибір автопарку, враховуючи показники продуктивності, погодні умови та маршрути транспортних доріг із точністю 85%. Бек та ін. використовували алгоритм глибокої нейронної мережі для прогнозування видобутку руди на шахті в Республіці Корея на основі показників продуктивності парку вантажівок [ 17 ]. Алгоритм машинного навчання використовувався для двох наборів даних: один для виробництва вранці, а другий – удень. Результати алгоритму глибокої нейронної мережі були багатообіцяючими, оскільки середня абсолютна відсоткова похибка для ранкового виробництва становила 11,40%, а для післяобіднього – 8,87% [ 17] . Середня абсолютна відсоткова помилка менше 10 є відмінною, тоді як від 10 до 25 є низькою, але прийнятною [ 18 ]. Середня абсолютна похибка порівнює прогнозований результат із відомим значенням, а на основі їх різниці розраховується відсоткова похибка [ 18 ]. Алгоритми та результати машинного навчання мають продовжувати вдосконалюватися з часом, і з додаванням збору великих даних у гірничодобувній промисловості вони, як очікується, використовуватимуться для прогнозування показників продуктивності та необхідного обслуговування.

## 3. Методика досліджень

### 3.1. Вихідні дані

У цьому дослідженні використовувалися дані двох різних програм: CAT Productivity і CAT MineStar Edge. Це програмне забезпечення є власністю Caterpillar і вимагає від окремих компаній оплатити підписку на підключення своїх машин до мережі для доступу до даних, зібраних датчиками на машині. Ці машини підключаються до мережі за допомогою сигналу 4G, який передається через так званий блок Product Link. Це вікно Product Link дозволяє передавати інформацію про працездатність, використання, виробництво та години/місце розташування на інформаційні панелі програмного забезпечення для використання клієнтом або власником програмного забезпечення [ 19 , 20 , 21]. У цьому дослідженні всі використовувані машини мають блок Product Link покоління PLE641, який дозволяє збирати розширені виробничі показники. У CAT Productivity було кілька завантажувачів, окрім тих, які використовувалися в цьому дослідженні, які мали коробки PLE641, але оскільки клієнти не підписалися на програмне забезпечення, ці машини відображали лише дані про використання, а також показники часу циклу. Ці дві частини програмного забезпечення можуть надавати схожі показники продуктивності, працездатності машини та використання, але є кілька ключових відмінностей. CAT MineStar Edge дозволяє збирати дані в режимі реального часу та відтворювати записи про те, що роблять машини, і легко визначає, який самоскид завантажується фронтальним навантажувачем [ 22]. З іншого боку, CAT Productivity не забезпечує збирання даних у режимі реального часу та не надає такої глибокої інформації, як CAT MineStar Edge, оскільки це дешевший продукт за підпискою [ 23 ]. Ці дані не є загальнодоступними, і вони були надані компанією ТОВ з П Цепелін Україна для використання в цьому дослідженні. Усі дані, зібрані під час цього дослідницького

дослідження, були очищені від будь-якої ідентифікаційної інформації клієнта, а також номерів активів, щоб запобігти проникненню в інформацію про виробництво клієнтів.

### 3.1.1. Набір даних продуктивності машин

Перші дані, зібрані з CAT Productivity, склалися зі змінних для базової виробничої інформації з набором даних для кожного із семи навантажувачів. Перша група даних використовувала дані за вісім місяців з квітня по грудень 2022 року. Навантажувачі, які використовувалися в цьому дослідженні, працювали в різних кар'єрах з подібними умовами виробництва. Ці кам'яні кар'єри є наземними підприємствами, де видобувають вапняк, пісок і гравій. Перша група даних була зібрана за допомогою програмного забезпечення Cat Productivity. Навантажувачі Caterpillar мали дещо різні розміри залежно від свого покоління. Сім навантажувачів Caterpillar склалися з однієї машини 992К, двох машин 990К, трьох машин 988К і однієї машини 988К ХЕ. Ці машини мають різні розміри, але часто використовуються для кар'єрів подібного розміру на вибір компанії. Дані навантажувача були прив'язані до кожної кар'єрної вантажівки, що дозволяло також збирати показники вантажівки. Зібрані дані змінні для дати й часу, корисного навантаження ковша (в тоннах), загальної кількості ковшів і загального корисного навантаження на вантажівку (в тоннах), тривалості циклу (у секундах) і ідентифікатора вантажівки. Корисне навантаження ковша – це лише тоннаж матеріалу, який знаходиться у ковші навантажувача до того, як його буде вивантажено в кар'єрний вантажівку. Загальна кількість ковшів вантажівки – це загальна кількість ковшів навантажувача, наповнених матеріалом, яка потрібна для заповнення вантажівки перед тим, як вона вирушить із зони завантаження. Загальне корисне навантаження вантажівки — це загальна тоннаж матеріалу, з яким

вантажівка відправляється із зони завантаження після того, як його завантажив навантажувач. Тривалість циклу стосується навантажувача та вимірюється загальним часом наступних чотирьох фаз: збирання матеріалу з робочого вибою для заповнення ковша навантажувача,24 ]. Ідентифікатор вантажівки — це просто спосіб для компаній визначити, яку вантажівку завантажує навантажувач. Таблиця 1 показує вибірку зібраних даних.

Таблиця 1.

Приклад виробничих даних Cat Productivity для навантажувачів із кам'яних кар'єрів.

Дата вимірювання	Маса в ковші, т	Кількість ковшів, шт	Час циклу, с	Планова вантажність, т	ID вантажівки
16 квітня 2024	16,94	4	68,93	45	1
12 травня 2024	14,56	5	87,64	73	1

### 3.1.2. Набір даних CAT MineStar Edge

Другий набір даних, що складається з двох завантажувачів від CAT MineStar Edge, мав багато змінних, але лише деякі з них можна використовувати для аналізу показників продуктивності. Це пов'язано з тим, що більша частина інформації в цьому наборі даних складається з конфіденційної для компанії інформації, як-от місцезнаходження завантаження та скидання в X, Y та Z, а також координати широти та довготи, маршрути перевезення з відмітною інформацією та серійний номер машини чисел. Окрім конфіденційної інформації, була різна інформація, яка була пропущена через те, що вона не була пов'язана з жодними показниками ефективності. Наприклад, тип змінної циклу було пропущено, оскільки кожна точка даних мала однакове значення, яке було «HAUL», а також інші змінні, такі як оператор перевезення, який не було заповнено для жодної точки даних. Другу групу склали дві машини-навантажувачі покоління Caterpillar 993K. Цей другий набір даних також складався з даних за вісім

місяців з великомасштабної вугільної шахти. Ці Caterpillar 993K є великими одиницями техніки, які здатні виконувати великі виробничі операції, такі як поверхня вугільної шахти, на якій проводилося це дослідження. Найкориснішими змінними в наборі даних були дата, тривалість навантаження, час циклу вантажівки (секунди), повна планова відстань (м), кількість ковшів навантажувача та дозволене корисне навантаження (тони). Тривалість завантаження — це тривалість часу, який знадобився цим навантажувачам, щоб заповнити кар’єрні вантажівки матеріалом. Час циклу вантажівки – це загальний час, протягом якого кар’єрна вантажівка проходить фази початку завантаження, транспортування до місця звалища, скидання, транспортування назад до місця завантаження,25 ]. Повна планова відстань — це загальна горизонтальна відстань, яку проходить кар’єрний вантажівка, щоб дістатися до місця звалища від місця завантаження. Кількість ковшів навантажувача — це кількість ковшів навантажувача, необхідних для заповнення платформи кар’єрної вантажівки перед тим, як вона відправиться на місце звалища [ 26 ]. Розширене корисне навантаження — інша назва загального корисного навантаження вантажівки. Таблиця 2 показує вибірку даних.

Таблиця 2.

Зразок виробничих даних Cat MineStar Edge для навантажувачів із відкритої вугільної шахти.

Дата вимірювання	Маса в ковші, т	Кількість ковшів, шт	Час циклу, с	Планова вантажність, т
16 квітня 2024	16,94	4	68,93	45
12 травня 2024	14,56	5	87,64	73

### 3.2. Попередня обробка даних

Кожен набір даних, завантажений з кожного завантажувача, очищався для аналізу даних, який потрібно було провести та ввести в моделі. Цей процес включав визначення діапазону тривалості циклу для кожного навантажувача (щоб вони не були ні надто великими, ні надто малими, вказуючи на завдання, окрім завантаження кар'єрних вантажівок), видалення відсутніх даних і фільтрування непотрібних змінних, які не вказували б на показники виробництва. Зважаючи на відсутність заповнених даних для деяких виробничих змінних, їх було видалено з наборів даних через неузгодженість у подальшому зборі даних для кожного завантажувача. Наприклад, дані, які не вказували на виробничі показники, являли собою таку інформацію, як координати місць завантаження та звалища або назви транспортних маршрутів, які ідентифікувала компанія.

Для першого набору даних від Cat Productivity потрібно було видалити численні точки даних. Машини підключаються до програмного забезпечення через пристрої, які називаються Product Link box. Блоки Product Link мають кілька поколінь, які збирають різні обсяги даних залежно від використовуваного покоління. Наприклад, багато завантажувачів використовували коробку Product Link покоління 541, тоді як два завантажувачі використовували коробки 641 Product Link. Враховуючи змішане покоління блоків Product Link серед завантажувачів, деякі змінні було видалено з набору даних через неузгодженість у подальшому зборі даних у кожному завантажувачі. Ці змінні включали послідовність корисного навантаження ковша, загальну кількість ковшів вантажівки, загальне корисне навантаження вантажівки та транспортну одиницю. Навантажувачі з коробками 641 Product Link були налаштовані, щоб визначити, скільки циклів ковша потрібно для завантаження вантажівки,

Для набору даних Cat MineStar Edge єдині змінні, які були видалені, були пов'язані з зоною завантаження, координатами маршруту транспортування та нахилом маршруту транспортування. Час циклу кожного завантажувача аналізувався в розподілі, щоб визначити частоту виконання

завдання, створюючи низькі та верхні межі для граничного часу. Оператори навантажувача можуть виконувати кілька завдань, чекаючи, поки кар'єрні вантажівки розмістяться поруч із навантажувачем на виробці. На низькому рівні вони можуть чистити обличчя або підлогу, швидко збираючи та скидаючи матеріал. На високому рівні вони могли сидіти з відром, наповненим матеріалом, і чекати, поки під'їде вантажівка.

### 3.3. Аналіз даних

Після збору даних за вісім місяців для кожного навантажувача з CAT Productivity і CAT MineStar Edge було проведено очищення даних. Microsoft Excel використовувався для очищення та підготовки наборів даних для перевірки точності та надійності даних. Процес очищення даних передбачав видалення конфіденційної інформації з компаній і змінних, які постійно мали брак інформації. Після видалення даних було розраховано коефіцієнт заповнення для набору завантажувачів CAT Productivity, який було проведено шляхом ділення корисного навантаження кожного ковша на максимальне корисне навантаження, можливе для цієї машини. Через відсутність корисного навантаження ковша індивідуального циклу в наборі даних CAT MineStar Edge не вдалося розрахувати коефіцієнт заповнення для двох навантажувачів у вугільній шахті. Після процесу очищення даних функція PivotTable у Microsoft Excel використовувалася для розрахунку середнього корисного навантаження для кожного часу циклу. Щоб визначити найпоширеніший час циклу для кожного завантажувача, частота, з якою з'являється кожен цикл, була розрахована в кожному наборі даних. Це дає змогу зрозуміти розподіл корисного навантаження та допомогти визначити будь-які закономірності чи тенденції в даних. Далі середнє корисне навантаження та частоту циклів імпортували в програмне забезпечення Google Colab (версія 3.9, Google, Маунтін-В'ю, Каліфорнія, США). Використовуючи це програмне забезпечення з бібліотеками Python, Matplotlib (версія 3.7.1), Pandas (версія 1.5.3) і NumPy (версія 1.22.4), були



створені діаграми розсіювання для відображення розподілу даних для середнього корисного навантаження та циклу час, і гістограми були створені для візуалізації частоти часу циклу для кожного із завантажувачів.

### 3.4. Статистичний аналіз

Google Colab використовувався для статистичного аналізу, обчислень і візуалізації даних для всіх наборів даних. Першим кроком було розрахувати показники центральної тенденції для обраних основних виробничих показників для кожного набору навантажувачів. Показники центральної тенденції включали середнє значення, стандартне відхилення та перший і третій квантиль. Для навантажувачів CAT Productivity порівнювали сім одиниць навантажувача в цьому наборі на основі змінних часу циклу навантажувача, коефіцієнта заповнення та корисного навантаження ковша. Для навантажувачів CAT MineStar Edge порівнювали два блоки навантажувача в цьому наборі даних на основі змінних загального корисного навантаження вантажівки, часу циклу навантажувача та часу циклу вантажівки. Через обчислення показників центральної тенденції було отримано візуалізацію розподілу даних, а також будь-які тенденції чи закономірності, які можуть з'явитися. Вони можуть надати уявлення про продуктивність кожного завантажувача в порівнянні один з одним, а також показати будь-які сфери вдосконалення. Функція щільності ймовірності (PDF) обчислюється за допомогою цих показників центральної тенденції, щоб візуалізувати розподіл даних. Завдяки візуалізації PDF-файлу для кожного завантажувача для кожної змінної можна визначити будь-які значні відмінності між продуктивністю завантажувачів у кожному наборі, щоб отримати краще розуміння основних даних. Не вистачає значущих виробничих змінних для порівняння даних, тому це обмеження. Було зроблено необхідні судження щодо того, на яких змінних зосередитися для порівняння виробничих показників. Функція щільності ймовірності (PDF) обчислюється за допомогою цих показників центральної тенденції, щоб

візуалізувати розподіл даних. Завдяки візуалізації PDF-файлу для кожного завантажувача для кожної змінної можна визначити будь-які значні відмінності між продуктивністю завантажувачів у кожному наборі, щоб отримати краще розуміння основних даних. Не вистачає значущих виробничих змінних для порівняння даних, тому це обмеження. Було зроблено необхідні судження щодо того, на яких змінних зосередитися для порівняння виробничих показників. Функція щільності ймовірності (PDF) обчислюється за допомогою цих показників центральної тенденції, щоб візуалізувати розподіл даних. Завдяки візуалізації PDF-файлу для кожного завантажувача для кожної змінної можна визначити будь-які значні відмінності між продуктивністю завантажувачів у кожному наборі, щоб отримати краще розуміння основних даних. Не вистачає значущих виробничих змінних для порівняння даних, тому це обмеження. Було зроблено необхідні судження щодо того, на яких змінних зосередитися для порівняння виробничих показників. Будь-які суттєві відмінності між продуктивністю завантажувачів у кожному наборі можна визначити, щоб отримати краще розуміння основних даних. Не вистачає значущих виробничих змінних для порівняння даних, тому це обмеження. Було зроблено необхідні судження щодо того, на яких змінних зосередитися для порівняння виробничих показників. Будь-які суттєві відмінності між продуктивністю завантажувачів у кожному наборі можна визначити, щоб отримати краще розуміння основних даних. Не вистачає значущих виробничих змінних для порівняння даних, тому це обмеження. Було зроблено необхідні судження щодо того, на яких змінних зосередитися для порівняння виробничих показників.

### 3.5. Визначення достовірності результатів

Аналіз обслуговування GET проводився за допомогою Microsoft Excel і Google Colab. Excel в основному використовувався для збору середніх значень виробництва та аналізу витрат. Спочатку було об'єднано два набори

даних: набір даних виробничих показників для завантажувача 988К, для якого проводилося технічне обслуговування, і набір даних про дні технічного обслуговування для заміни ресурсу GET. Спочатку середнє значення корисного навантаження для кожного дня обчислювалося за допомогою функції PivotTable. Ці дані щодо середнього корисного навантаження за кожен день потім зіставлялися з днями, коли проводилося технічне обслуговування. Потім було розраховано середнє значення корисного навантаження за період у п'ять днів до та після технічного обслуговування, чотири дні до та після, три дні до та після, два дні до та після, один день до та після та в день технічного обслуговування. Це дало таблицю з 11 значень, які показували середню продуктивність до та після кожного дня технічного обслуговування. Потім Google Colab використовувався з бібліотекою matplotlib.pyplot у Python для створення візуалізації розподілу цих середніх значень корисного навантаження навколо технічного обслуговування. Потім було проведено вартісний аналіз вартості продукції за ці періоди. Це було досягнуто за допомогою деяких припущень, які перелічені в Таблиця 3. Для спекулятивного аналізу вапнякової породи на той час була обрана ціна 5 доларів США за тону продукту.

Таблиця 3.

Припущені значення для розрахунку вартості виробництва для кожного набору днів виробництва до та після технічного обслуговування.

Середня ціна Продукт (USD/тонна)	Середній Час технічного обслуговування (година)	Час виготовлення На день (годин)	Середня кількість ковшів на вантажівку	Середній цикл Час (секунди)	Ціна зубів (USD)
5	1	7,5	7	50	2181,76

Після того, як були розраховані значення виробництва для кожного періоду до і після технічного обслуговування, потім було розраховано відсоткову зміну шляхом порівняння значень до і після кожного дня технічного обслуговування. Ці відсоткові зміни та дні технічного

обслуговування потім були введені в Google Colab, і знову, за допомогою бібліотеки `matplotlib.pyplot`, вони були нанесені на графік для візуалізації даних.

Налаштування та аналіз машинного навчання проводилися за допомогою Google Colab, оскільки він має доступ до бібліотек машинного навчання в Python. Спочатку всі дані завантажувача були об'єднані в один набір даних. Використовуючи Google Colab, вони спочатку були введені в алгоритм лінійної регресії для визначення кореляції між даними в наборі даних. R-квадрат використовується як показник продуктивності моделі з цими моделями регресії. Через погані результати було створено іншу модель лінійної регресії з використанням лише випадкової частини набору даних, щоб перевірити, чи є чіткіші результати. Оскільки лінійна регресія дала погані результати, для візуалізації кореляції між значеннями в даних було обрано поліноміальну регресію. Дані були проштовхнуті через цю модель поліноміальної регресії, і отримано R-квадрат. Коли це було завершено, теплова карта була створена для візуалізації кореляції між змінними в даних. Наступним кроком було створення моделей для алгоритмів прогнозування машинного навчання. Чотири моделі, використані в цьому аналізі, були алгоритмами прогнозування на основі K-найближчих сусідів, поліноміальної регресії, дерева рішень і випадкового лісу. Вхідні дані в ці моделі базувалися на наборах для навчання та тестування. Приблизно 20% даних було обрано як набір для тестування, а 80% – як набір даних для навчання. Оцінка точності є результатом цих моделей, який вказує на продуктивність моделі. Чотири моделі, використані в цьому аналізі, були алгоритмами прогнозування на основі K-найближчих сусідів, поліноміальної регресії, дерева рішень і випадкового лісу. Вхідні дані в ці моделі базувалися на наборах для навчання та тестування. Приблизно 20% даних було обрано як набір для тестування, а 80% – як набір даних для навчання. Оцінка точності є результатом цих моделей, який вказує на продуктивність моделі. Чотири моделі, використані в цьому аналізі, були алгоритмами прогнозування на

основі  $K$ -найближчих сусідів, поліноміальної регресії, дерева рішень і випадкового лісу. Вхідні дані в ці моделі базувалися на наборах для навчання та тестування. Приблизно 20% даних було обрано як набір для тестування, а 80% – як набір даних для навчання. Оцінка точності є результатом цих моделей, який вказує на продуктивність моделі.

## 4. Результати та їх аналіз

### 4.1. Аналіз даних

Завдяки аналізу даних кожного набору даних завантажувача було отримано розподіл середнього корисного навантаження для кожного часу циклу. Ці дані були розділені, щоб визначити, чи існує оптимальна точка для кожного навантажувача, при якій середнє корисне навантаження на ковш можна збільшити за швидшого циклу, щоб отримати більшу продуктивність на лінії. Також було включено частоту кожного циклу, щоб визначити розподіл даних і те, як часто оператори досягають кожного циклу. Кожен завантажувач аналізувався між циклами, які були визначальними для його розподілу даних. Тривалість циклу семи навантажувачів, які працюють у каменоломнях, становить приблизно від 25 до 80 с. На рис. 1-7 показані діаграми розсіювання середнього корисного навантаження в залежності від часу циклу та гістограми частоти кожного часу циклу.

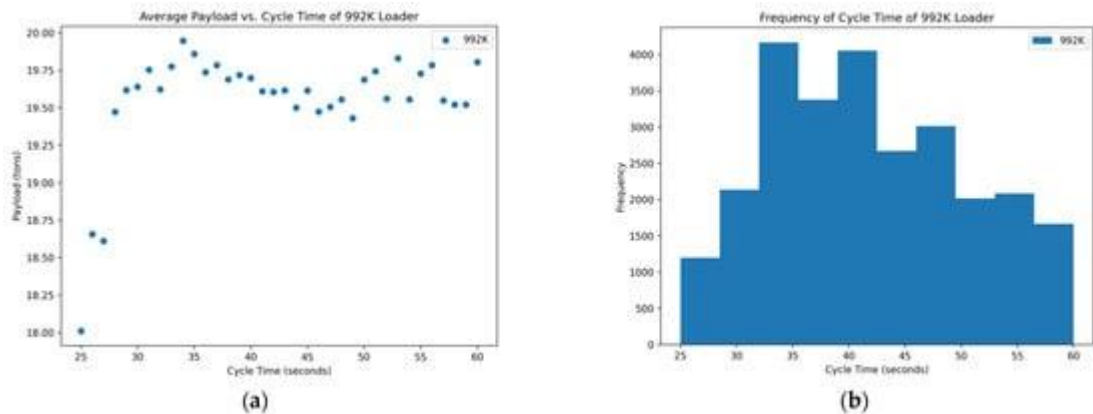
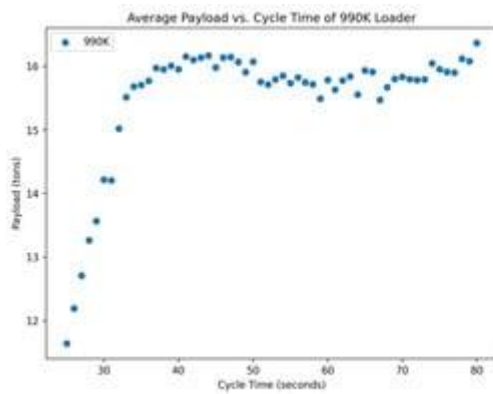
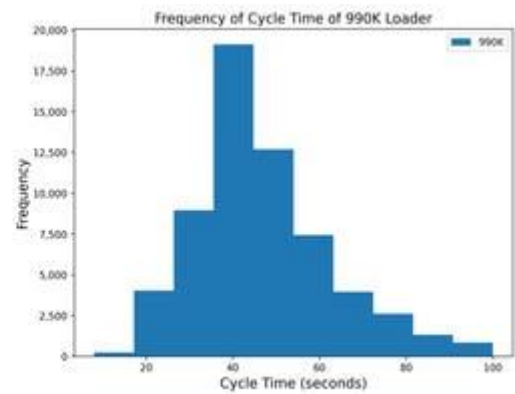


Рисунок 1. Розподіл змінних у наборі даних 992К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.

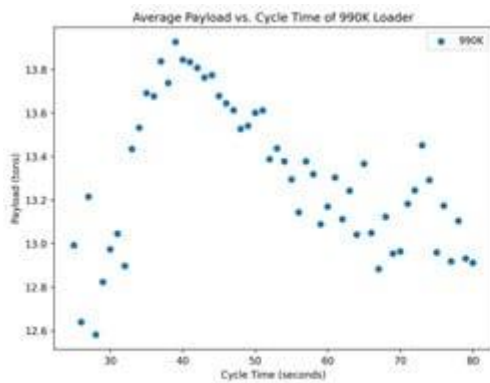


(a)

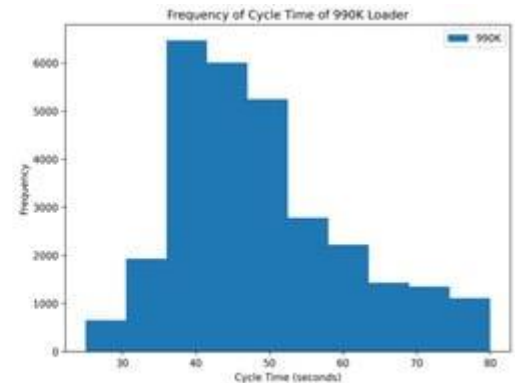


(b)

Рисунок 2. Розподіл змінних у наборі даних 990К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.

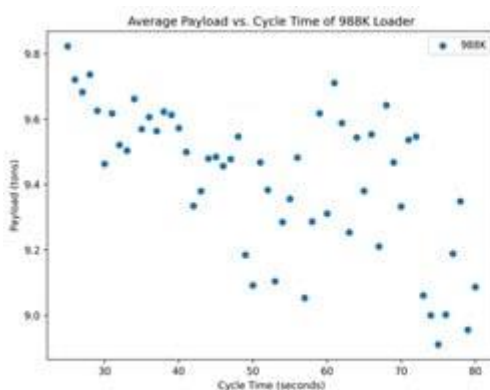


(a)

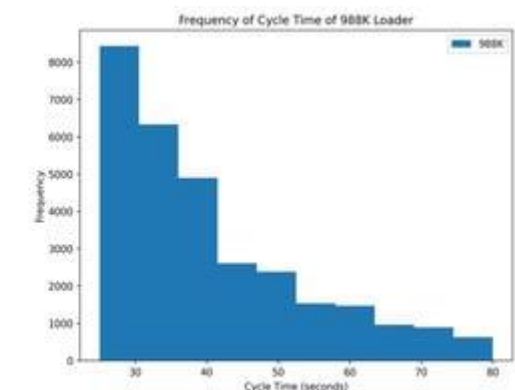


(b)

Рисунок 3. Розподіл змінних у наборі даних 990К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.

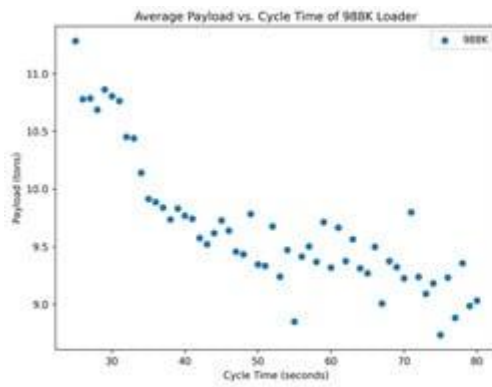


(a)

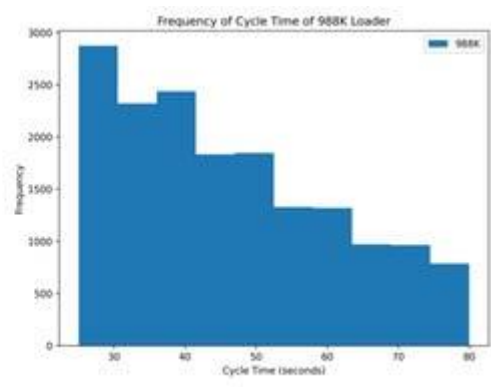


(b)

Рисунок 4. Розподіл змінних у наборі даних 988К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.



(a)

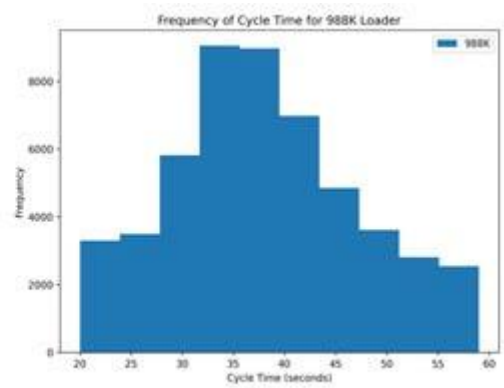


(b)

Рисунок 5. Розподіл змінних у наборі даних 988К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.



(a)

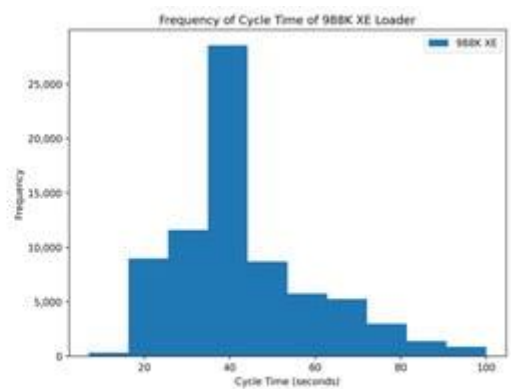


(b)

Рисунок 6. Розподіл змінних у наборі даних 988К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.



(a)



(b)

Рисунок 7. Розподіл змінних у наборі даних 988К ХЕ: ( а ) діаграма розсіювання середнього корисного навантаження на сегмент і час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.



З даних, які були зібрані та відображені для семи навантажувачів у каменоломнях, три з них показали багатообіцяючі результати, два показали непереконливі і два показали погані результати щодо тривалості циклу та продуктивності. Діаграми розсіювання на рисунку 3, рисунку 6 та рисунку 7 демонструють багатообіцяючі результати. Ці цифри вказують на оптимальну точку навколо часу циклу приблизно 35–40 с, у якому корисне навантаження призводить до вищого виходу. Тенденції діаграм розсіювання ілюструють збільшення продуктивності корисного навантаження, коли вони наближаються до верхньої межі часу циклу. Збільшення корисного навантаження теоретично пояснюється тим фактом, що під час довшого циклу оператори неодноразово закопують лопатами підірвані камені, щоб отримати вищий коефіцієнт наповнення своїх ковшів. Діаграми розсіювання на рис. 1 і 2 показують непереконливі результати. Діаграми на рисунках 1 і 2 є непереконливими, тому що, незважаючи на те, що існує оптимальна точка для часу циклу, який усереднює високий вихід корисного навантаження, він вирівнюється після цієї точки і має лише незначні зміни середнього корисного навантаження в кожному циклі. час. Оптимальною точкою для цих показників також є приблизно 35–40 с. На рисунках 4 і 5 показано погані результати. Рисунок 4 і Рисунок 5 мають негативну тенденцію з багатьма викидами без чіткої оптимальної точки для високого виробництва, за винятком нижньої межі для часу циклу. Нижня межа часу циклу була розташована за межами загальноприйнятого розумного часу циклу, що могло свідчити про погані результати.

Гістограми частоти кожного часу циклу для кожного завантажувача в основному показують нормальний розподіл навколо часів циклу, які представляють оптимальну точку для високого виходу корисного навантаження. Кілька гістограм, а саме на рисунку 2, рисунку 3 та рисунку 7, зміщені вправо, що може свідчити про те, що оператори виконують багато завдань, а не просто завантажують вантажівки. Більше даних біля верхньої межі тривалості циклу може свідчити про те, що оператори витрачають

непотрібну кількість часу на завантаження вантажівки або очищення дна кар'єру, що може вплинути на виробництво. Нормальний розподіл на цих графіках свідчить про те, що оператори цих навантажувачів працюють винятково. Цей перекис означає, що в цей час циклу існує ще багато точок даних із високим корисним навантаженням, які можуть збільшити дані. Рисунок 4 і Рисунок 5 не показують нормального розподілу, а натомість сильно зміщені вправо. Правий перекис суперечить розумному часу циклу, оскільки вони були б надто швидкими в цих точках і не могли б розумно підтримувати високу продуктивність корисного вантажу та скидати його в вантажівку за той самий мінімальний проміжок часу. Рисунки 4 і 5 свідчать або про аномалії, або про помилки в даних, і потребують додаткового вивчення з відвідуванням місця для аналізу операцій.

Два навантажувачі з відкритої вугільної шахти були проаналізовані так само, як і сім навантажувачів з кар'єрів. Однак ці навантажувачі не мали індивідуального часу циклу для кожного корисного навантаження ковша у своєму наборі даних. Як наслідок, на рисунках 8 і 9 показано середнє корисне навантаження на вантажівку в порівнянні з часом циклу заповнення цієї вантажівки, а також частотою кожного циклу, зазначеного поруч.

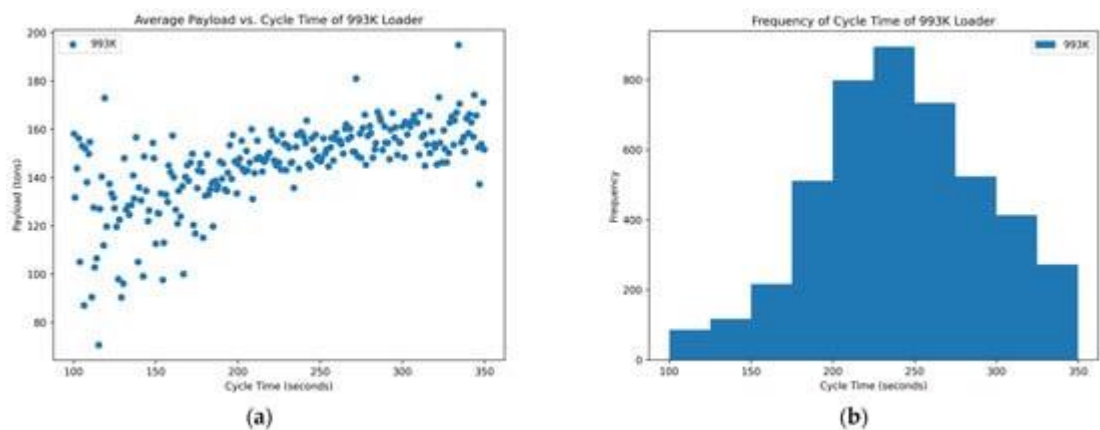


Рисунок 8. Розподіл змінних у першому наборі даних 993К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.

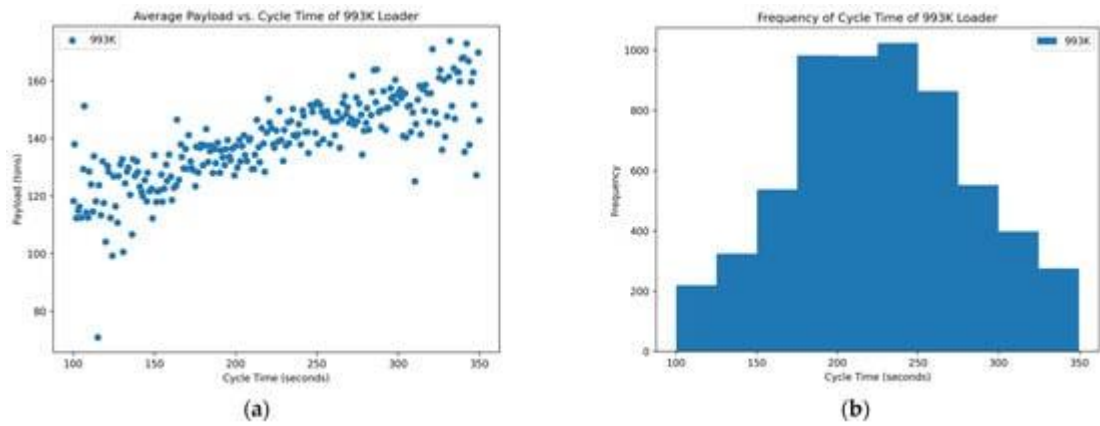


Рисунок 9. Розподіл змінних у другому наборі даних 993К: ( а ) діаграма розсіювання середнього корисного навантаження на відро та час циклу та ( б ) гістограма частоти кожного часу циклу, що виникає в наборі даних.

Два завантажувачі 993К конкретно не містять оптимальної точки навколо часу циклу, який демонструє високу ефективність корисного навантаження. Хоча діаграми розсіювання не показують оптимальної точки, вони мають позитивну тенденцію з деякими викидами. Позитивна тенденція підтверджує гіпотезу, згідно з якою ви збільшуєте час циклу, ви можете мати більший вихід корисного навантаження через неодноразове загрибання підірваних каменів для досягнення вищого коефіцієнта заповнення. Незважаючи на те, що це правда, тривалішу стратегію завантаження слід пом'якшити, оскільки тоді може виникнути ненавмисний ефект, коли завантаження однієї кар'єрної вантажівки протягом більшої кількості часу може підтримати всю операцію. Низька продуктивність за зміну може бути пов'язана з більшою тривалістю циклу або резервним копіюванням через тривалий час у черзі на місці завантаження.

Гістограми двох завантажувачів 993К на відкритих вугільних шахтах передусім показують нормальний розподіл, зосереджений навколо часу циклу 250 с. Ця центральна точка може свідчити про те, що ці більші операції, такі як ці вугільні шахти, ідеально працюють приблизно в цей час, оскільки вони частіше потрапляють на них. Нормальний розподіл із широким

діапазоном також свідчить про те, що ці вантажні одиниці використовуються не лише для завантаження кар'єрних вантажівок. Їх можна використовувати для очищення підлоги, а також для переміщення матеріалу зі складів у дробарки для продовження виробництва.

## 4.2. Статистичний аналіз

Було проведено статистичний аналіз певних змінних у даних, щоб візуалізувати їх розподіл для порівняння. Тривалість циклу, коефіцієнт заповнення, корисне навантаження, корисне навантаження вантажівки та час циклу вантажівки були визначені як найкращі для визначення виробничих характеристик навантажувачів у їхній відповідній операції. У цьому аналізі дані знову були розділені на два набори на основі типу кар'єру. Один набір даних включав кам'яні кар'єри, що склалися з семи навантажувачів, які обговорювалися в розділі збору даних. Інший набір даних включав два навантажувачі у відкритих вугільних шахтах, які також обговорювалися в наборі збору даних. Змінними для першого набору, які аналізувалися, були час циклу, коефіцієнт заповнення та корисне навантаження (рис. 10, рис. 11 і рис. 12).). Змінні, представлені на рисунку 10, рисунку 11 і рисунку 12, були обрані тому, що всі їхні точки даних були заповнені для семи завантажувачів у наборі даних.

Нормальний розподіл тривалості циклу для семи навантажувачів загалом був схожий один на одного. Розподіли були зосереджені навколо різних значень, але численні розподіли мали однакові розкиди розмірів. Кілька розподілів були зосереджені навколо часу циклу приблизно 50 с, що свідчить про те, що це середній час циклу для кар'єрів такого розміру. Два 988К мали найкращий середній час циклу близько 39 с, що вказувало на те, що вони були найшвидшими завантажувачами, але дивлячись на Рисунок 12 показує, що вони не мали великих значень корисного навантаження. Нормальний розподіл коефіцієнтів заповнення для кожного з навантажувачів значно відрізнявся.

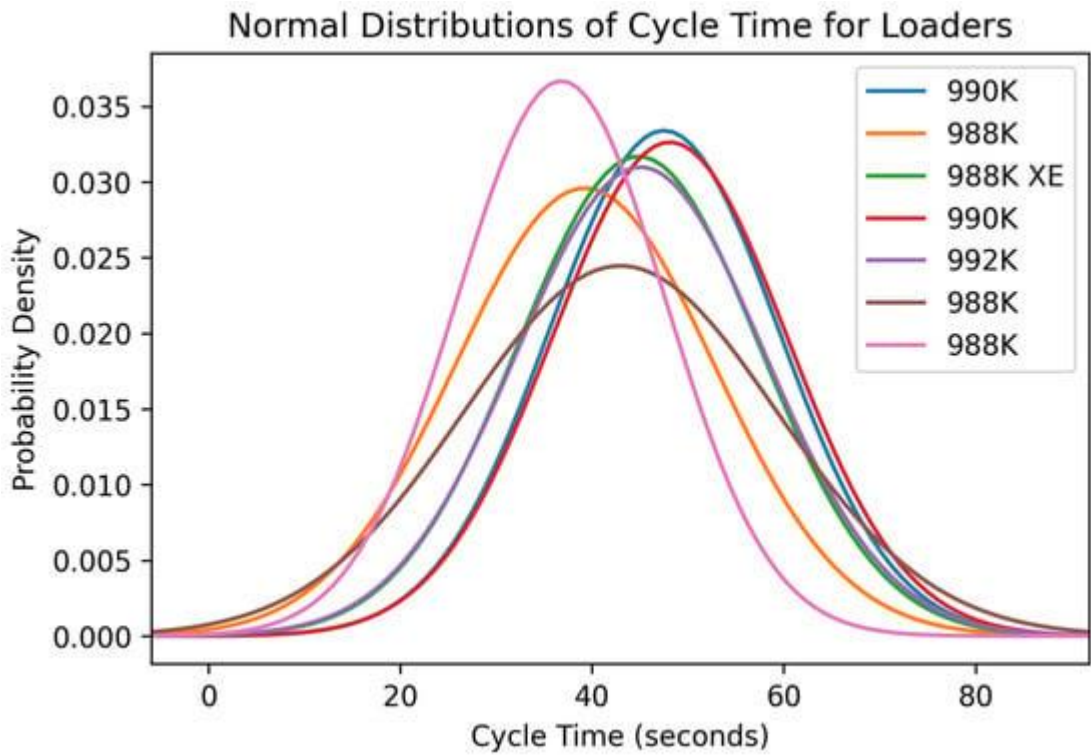


Рисунок 10. Нормальний розподіл змінного часу циклу для семи навантажувачів у кар'єрах.

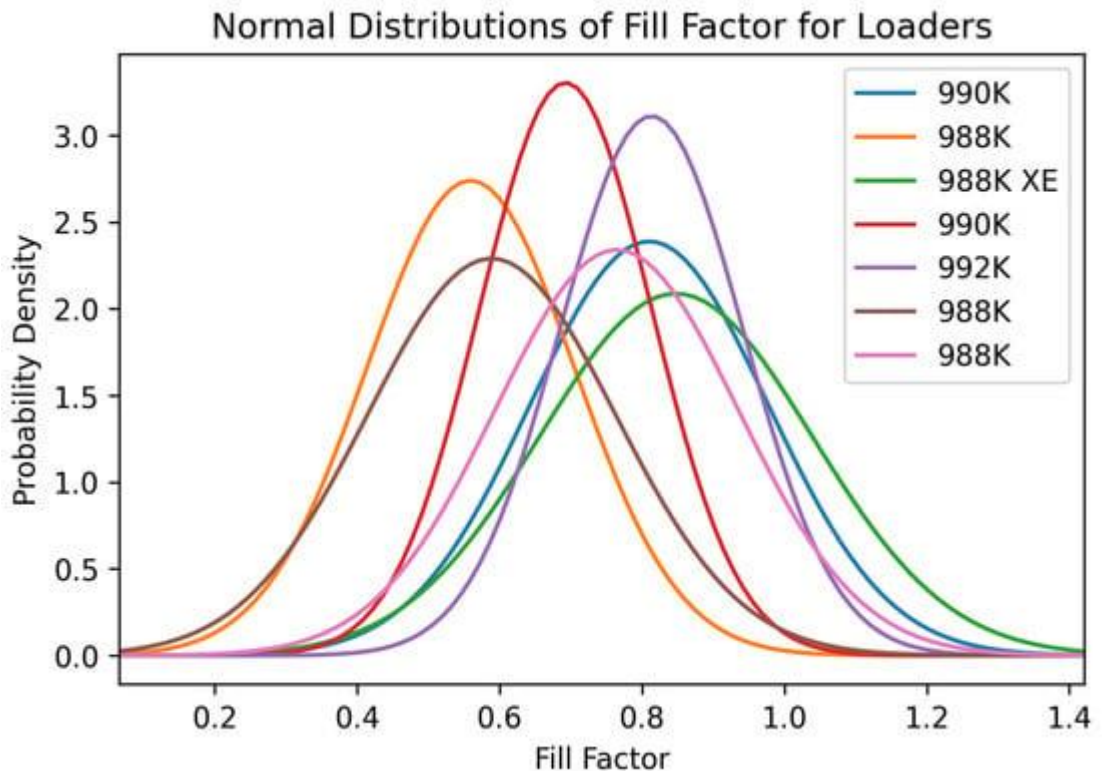


Рисунок 11. Нормальний розподіл змінного коефіцієнта заповнення для семи навантажувачів у кар'єрах.

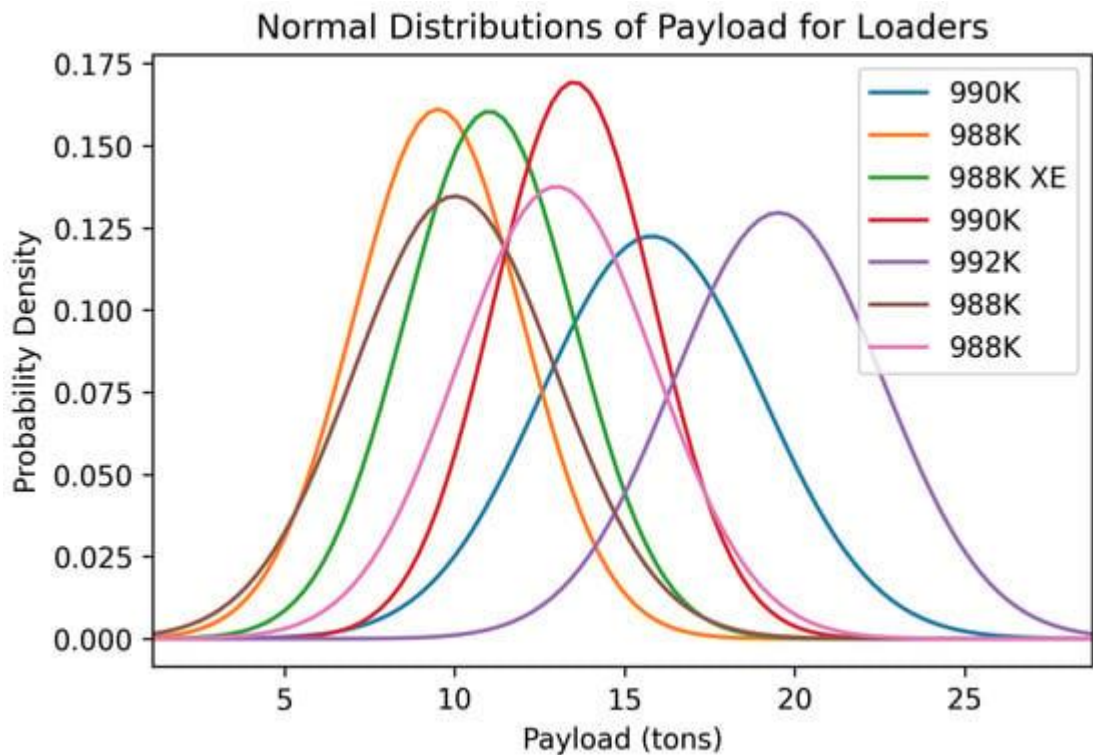


Рисунок 12. Нормальний розподіл змінного корисного навантаження для семи навантажувачів у кар'єрах.

Багато з них мали низькі коефіцієнти заповнення, що вказує на те, що оператори не використовують розмір своїх відер у повній мірі. Однією з причин недостатнього використання ковшів може бути відсутність досвіду оператора, а також завантажувальні пристрої, які часто використовуються для інших цілей, таких як очищення обличчя під час реєстрації цих завдань як циклу завантаження. Чотири завантажувачі, 992К, один 990К, один 988К і 998К ХЕ, показали оптимальні коефіцієнти заповнення приблизно 80% або вище. Високі коефіцієнти заповнення є ідеальними, оскільки оператори майже повністю використовують машини для заповнення вантажівок, що призводить до меншої кількості циклів на вантажівку та меншого часу циклу вантажівки. Нормальний розподіл корисного навантаження ковша кожної з одиниць завантаження значно відрізняється. Кожне покоління завантажувальних одиниць має різні розміри з ковшами різного розміру, які дозволяють використовувати більші або менші значення корисного

навантаження. Незважаючи на те, що розподіли демонструють різну потужність, кар'єри, у яких вони працюють, майже такого ж розміру, як аналогічний річний обсяг виробництва. Враховуючи схожі профілі кар'єрів і різні варіанти машин, навантажувач 992К був би найвигіднішим для цього типу роботи з відповідним оператором. Загалом 992К був одним із найефективніших навантажувачів за кожною зі змінних, показаних у цих розподілах. 992К мав вищий середній час циклу, ніж деякі інші навантажувачі, але залишався близьким до нижнього діапазону ідеального часу циклу для навантажувачів у цьому типі роботи. Загальне корисне навантаження вантажівки, час циклу навантажувача та час циклу вантажівки були проаналізовані в другому наборі даних відкритих вугільних шахт, що складається з двох навантажувачів. Набори даних вугільної шахти для вантажівки 993К містили більше даних, що дозволяло аналізувати змінні, найбільш пов'язані з показниками виробництва. На рисунках 13, 14 і 15 показано нормальні розподіли цих двох навантажувачів. Змінне корисне навантаження ковша було недоступне для цього набору даних, оскільки програмне забезпечення Cat MineStar Edge показує загальне корисне навантаження для кожної вантажівки замість корисного навантаження кожного циклу. Дві навантажувальні одиниці 993К є машинами того самого типу та розміру, але одна з одиниць перевершила іншу за двома з трьох змінних розподілів. Вивчаючи розподіл загального корисного навантаження вантажівки на рисунку 13, машина 993К, представлена синьою лінією, мала більшу загальну корисну навантаження вантажівки приблизно на 15 тонн. 993К, зображений помаранчевою лінією на рисунках 14 і 15, мав менший час циклу навантажувача та вантажівки з невеликим відривом. У майбутньому навантажувач 993К може продемонструвати вищу продуктивність, навіть якщо він трохи відстає за часом циклу. На змінні виробництва також може сильно впливати досвід роботи оператора, місце роботи в межах однієї шахти та тип матеріалу, над яким працюють вантажні установки в межах кар'єру.

Різні матеріали можуть бути важчими для копання або повністю розбитими під час вибухових робіт, що полегшує копання та впливає на виробництво.

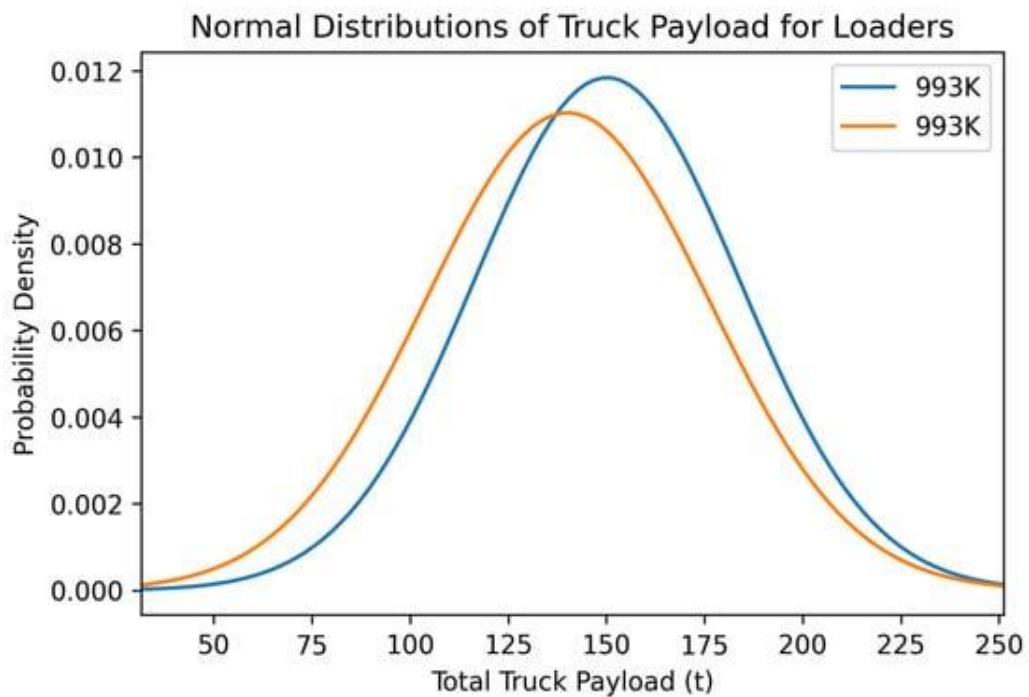


Рисунок 13. Нормальний розподіл змінного загального корисного навантаження вантажівки для кожного навантажувача у вугільній шахті.

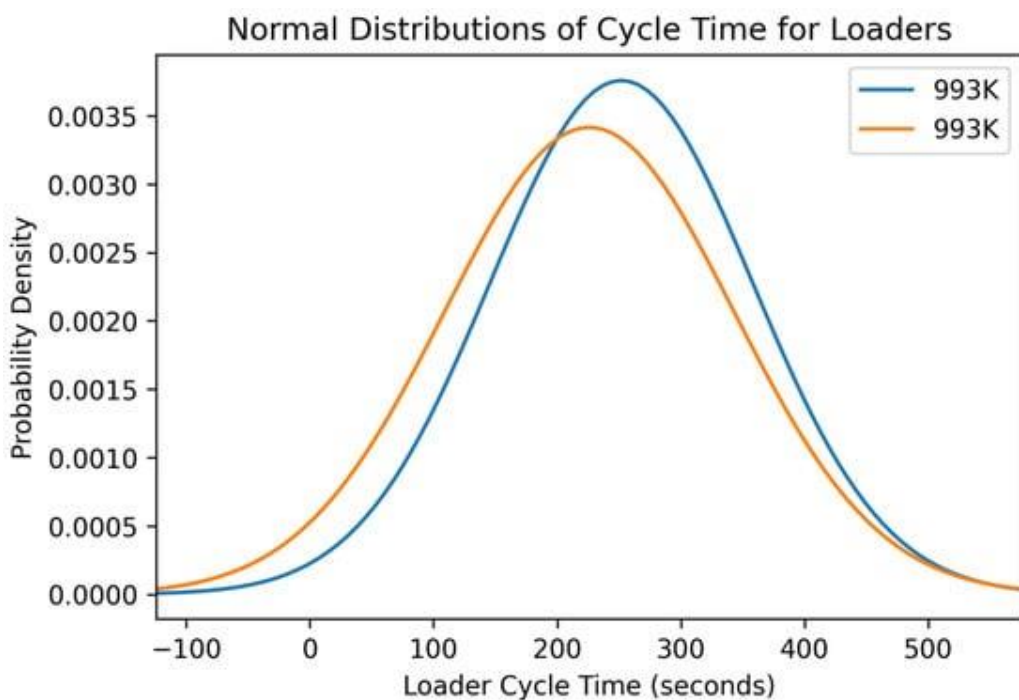


Рисунок 14. Нормальний розподіл змінного часу циклу завантажувача для кожного завантажувача у вугільній шахті.



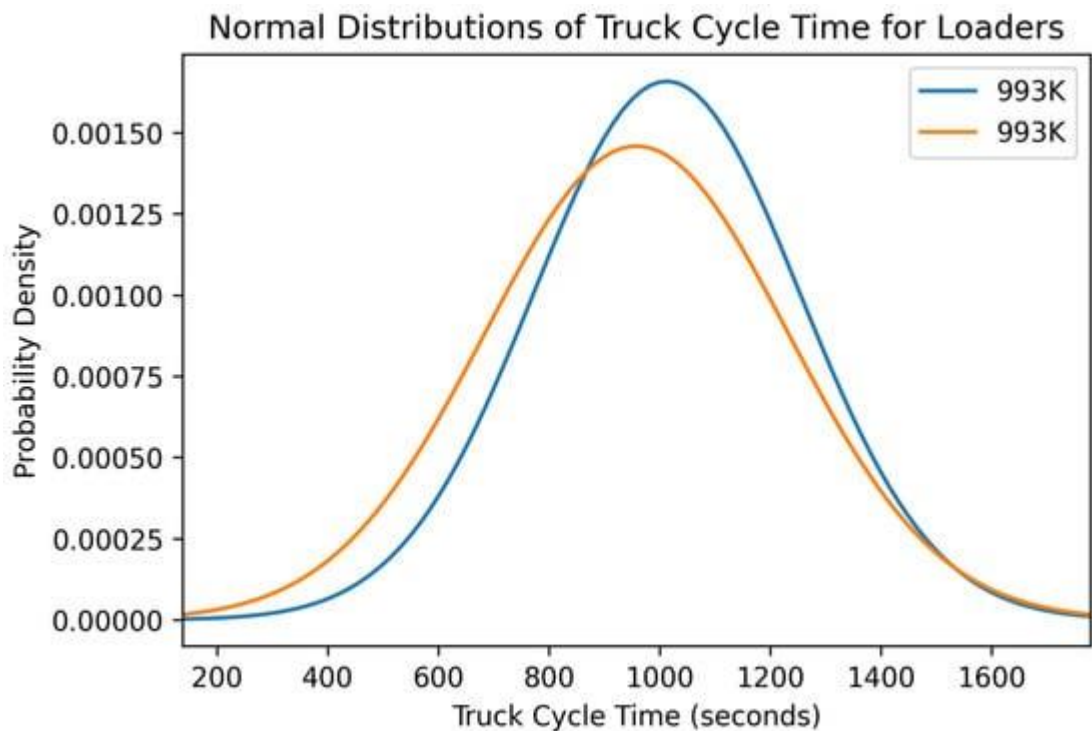


Рисунок 15. Нормальний розподіл змінного часу циклу вантажівки для кожного навантажувача у вугільній шахті.

Використовуючи значення продуктивності корисного навантаження з Cat Productivity і записи технічного обслуговування змін GET, було проведено аналіз змін у продуктивності до та після технічного обслуговування. Лише один запис про технічне обслуговування було отримано для однієї вантажної одиниці 988К з одного кар'єру. Цей запис технічного обслуговування містив лише дати технічного обслуговування. Це поєднувалося з виробничими показниками протягом усього часу технічного обслуговування з метою аналізу. Поточну практику технічного обслуговування визначає оператор машини. Якщо вони вважають, що зуби зношені на основі відчуття машини, тоді вони вимагають заміни. На жаль, багато операторів можуть не записувати, коли вони виконують технічне обслуговування, або використовувати паперові записи, які легко можна загубити, щоб визначити, коли відбулися операції з техобслуговування GET. Рисунок 16 показує середню продуктивність кожного вибору днів до технічного обслуговування. Вибір днів до технічного обслуговування

означає, що  $-4$  визначає чотири дні до технічного обслуговування,  $-2$  визначає два дні до технічного обслуговування,  $2$  визначає два дні після технічного обслуговування і так далі до п'яти днів до та після технічного обслуговування. На рисунку 17 показано процентну зміну вартості продукції. Дата технічного обслуговування базується на даті кожної зміни в GET, виконаної на машині 998К. Середнє корисне навантаження, що використовується в аналізі виробничих витрат, було визначено шляхом взяття набору днів від попереднього дня технічного обслуговування до відповідного дня, а потім також узявши набір днів від наступного дня технічного обслуговування до попереднього дня, про який йдеться. Відсоткова зміна була розрахована на основі цих днів.

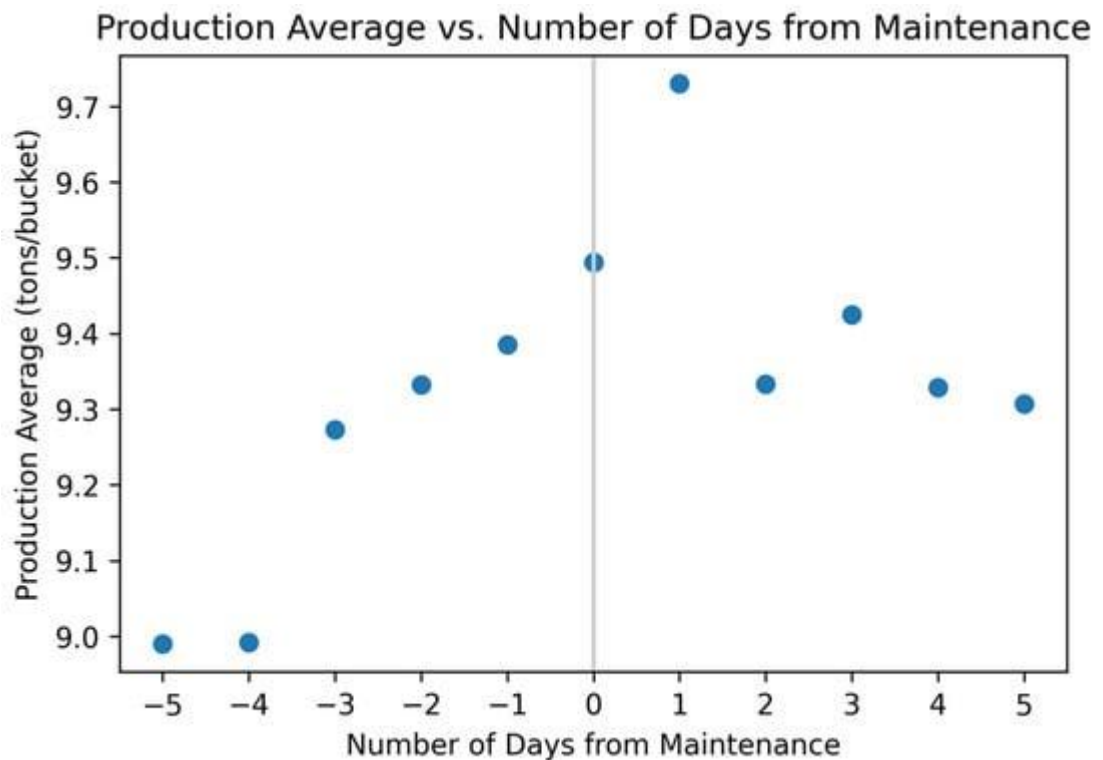


Рисунок 16. Для одного з навантажувачів 988К середня продуктивність для кожного набору днів до та після технічного обслуговування GET.

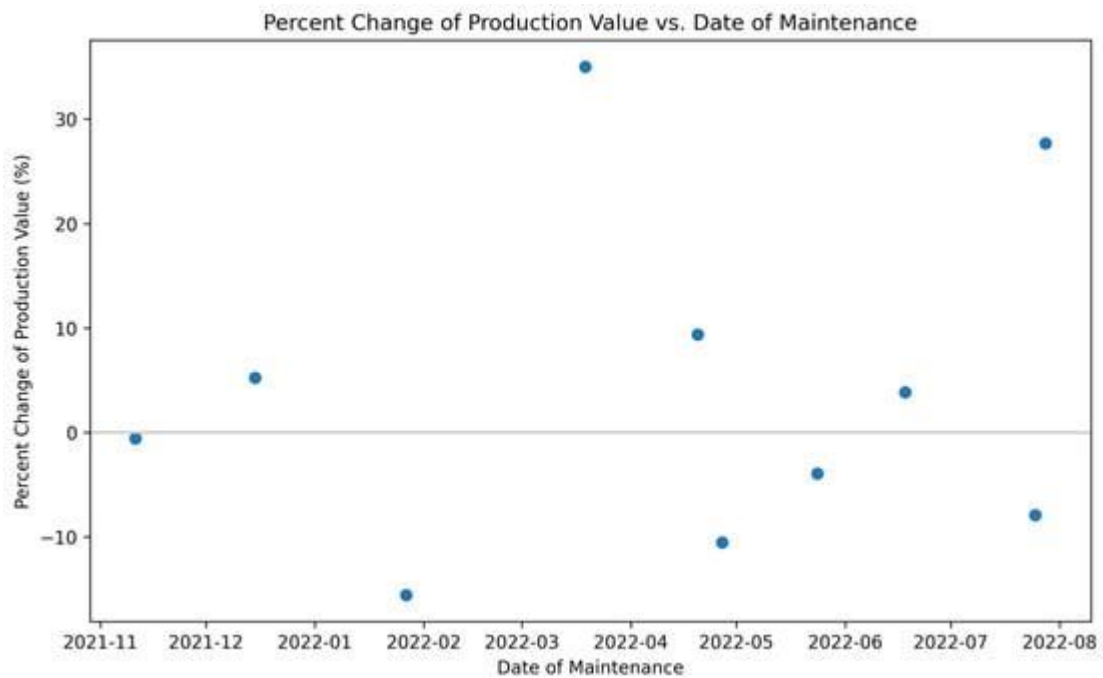


Рисунок 17. Для одного з навантажувачів 988К відсоткова зміна вартості виробництва до та після дати технічного обслуговування.

Середня продуктивність до і після технічного обслуговування представлена на рисунку 16 припускає, що після заміни GET на ковші навантажувача немає помітних змін у виробництві. Два набори днів виробництва, -4 і -5, мають низькі середні корисні навантаження, але при порівнянні наборів від -3 до -1 з наборами від 2 до 5, немає відносно ніяких змін у середньому виробництві. Виробництво зростає до технічного обслуговування, а в день технічного обслуговування воно є високим порівняно з іншими значеннями. Високе значення може бути пов'язане з виконанням технічного обслуговування на початку зміни, але в записах немає іншого індикатора, окрім дати, коли було виконано технічне обслуговування. Набір за один день після технічного обслуговування має найвищий середній показник продуктивності, але може вважатися викидом, якщо порівнювати його з наборами днів після технічного обслуговування.

Рисунок 17 ілюструє відсоткову зміну вартості виробництва в день технічного обслуговування та передбачає відсутність істотних змін у виробництві через технічне обслуговування. Відсоткова зміна охоплює рівномірний розподіл значень негативної та позитивної зміни. Варто

зазначити, що позитивна відсоткова зміна має вищий діапазон, що досягає приблизно 32%, тоді як негативна відсоткова зміна знижується лише до -15%. Спочатку припускали, що якщо змінити GET на ковші, це піде на користь операції та збільшить виробництво. Це буде результатом нових зубів, які дозволять легше копати. На основі даних і результатів, обчислених на основі цих даних, результати непереконливі щодо того, чи є поточний процес обслуговування GET корисним, чи оператори змінюють GET занадто часто. Технічне обслуговування GET, як правило, визначається рішенням оператора, оскільки їм важче вкопуватися в скелю або вони візуально помічають надмірний знос. Тому для кращих результатів у майбутньому має бути стандартний низький розмір зубів, який об'єктивно вказує на те, що необхідно виконати технічне обслуговування, щоб замінити їх.

### 4.3. Статистична обробка результатів

Використовуючи Python, для аналізу наборів даних використовувалися методи машинного навчання, щоб спочатку визначити кореляції між включеними змінними. Перший аналіз машинного навчання використовувався для тестування лінійної регресії. Оскільки корисне навантаження є ключовим показником продуктивності, воно було пов'язане з часом циклу. Використовуючи відповідні бібліотеки лінійної регресії в Python, було визначено тестовий і навчальний набір даних. Два набори потім використовувалися в алгоритмі лінійної регресії для перевірки змінних. Коефіцієнти кореляції були виведені для тестового набору після завершення навчання. Усі кореляції були меншими за 0,01 для кожного навантажувача, що вказує на те, що не було значної лінійної кореляції між вхідними змінними від навантажувача та самоскида.

Через погані результати, отримані за моделлю лінійної регресії, дані були проаналізовані за допомогою моделі поліноміальної регресії. Після введення даних для всіх завантажувачів і їх проходження через відповідну

бібліотеку поліноміальної регресії Sklearn на рисунку 18 показано теплову карту кореляції між змінними в наборі даних.

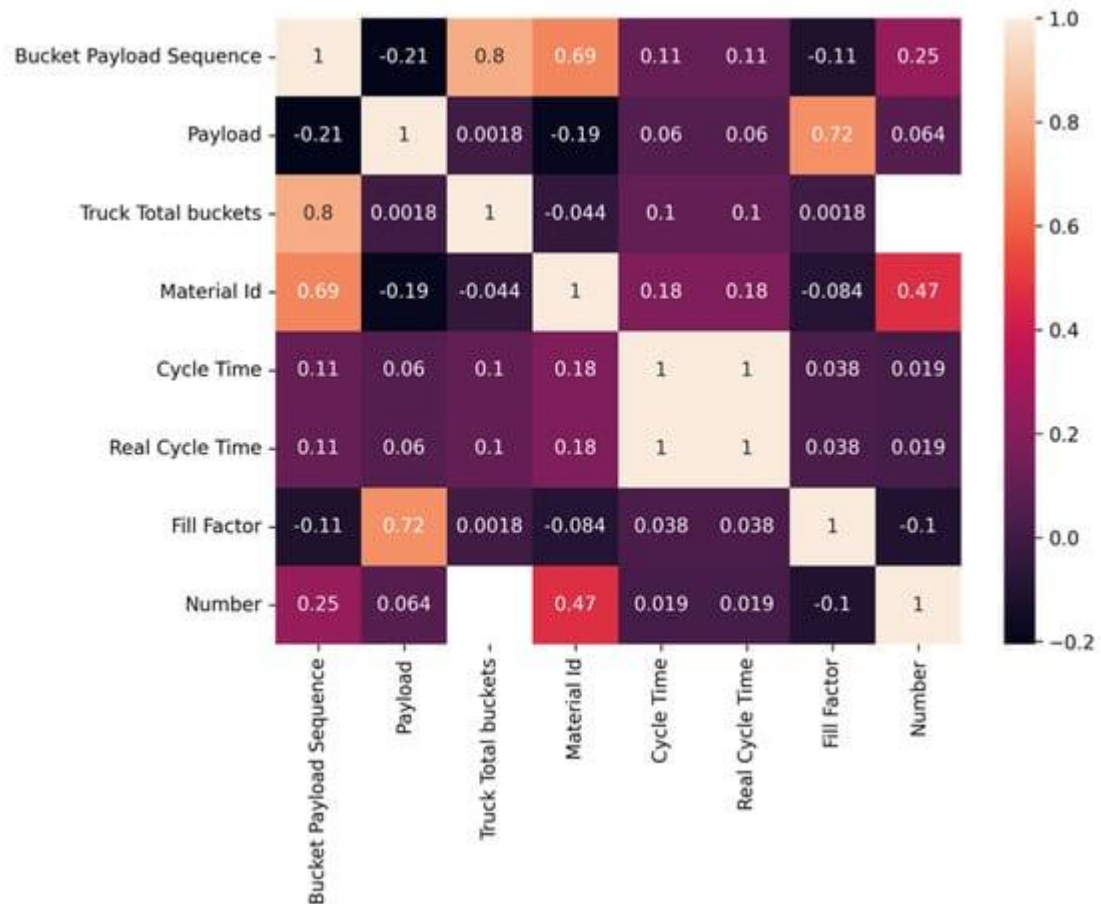


Рисунок 18. Теплова карта, створена за допомогою алгоритму машинного навчання поліноміальної регресії, щоб показати коефіцієнт кореляції між змінними в наборі даних завантажувача.

Результати алгоритму поліноміальної регресії не були обнадійливими. Незважаючи на більш високу кореляцію між декількома змінними порівняно з алгоритмом лінійної регресії, певна група змінних зі значущою кореляцією між ними не виявилася. Цей результат був несподіваним, оскільки очікувалося, що велика кількість даних у наборі даних принаймні приведе до певної форми кореляції між кількома змінними. Зокрема, очікується більш висока кореляція між змінними корисного навантаження та часу циклу. Результат кореляції між корисним навантаженням і часом циклу становив лише 0,06. Ми припустили, що буде показана значна кореляція, але з даними

та створеними моделями цей зв'язок не був знайдений. Через незвичайний характер результатів, майбутні дослідження потрібно буде проводити з додатковими типами машин і більше даних із заповненням усіх змінних. Це дослідження буде корисним для майбутнього покоління моделей.

Наступним методом машинного навчання були моделі прогнозування. У цьому дослідженні використовувалися моделі К-найближчих сусідів, дерева рішень, випадкові ліси та поліноміальна регресія. Повний набір даних було перевірено через кожен із цих алгоритмів класифікатора машинного навчання та виведено показник точності. Цей показник точності визначає, наскільки добре алгоритм може передбачати значення на основі навчального набору даних, який спочатку використовувався в алгоритмі. Таблиця 4 показує показники точності чотирьох моделей від найбільшого до найменшого.

Таблиця 4.

Оцінки точності чотирьох алгоритмів машинного навчання,  
використаних для набору даних для прогнозів.

Метод машинного навчання.	Точність передбачення
Класифікатор дерева рішень	0.8799
Випадковий ліс	0,8149
К-найближчі сусіди	0,766
Поліноміальна регресія	0,455

З чотирьох використаних алгоритмів класифікатор дерева рішень був найточнішим, за яким йшли алгоритми випадкового лісу та К-найближчих сусідів. З точністю 87,99% класифікована модель дерева рішень могла оцінити продуктивність будь-якого завантажувача на основі повного аналізу вхідних показників продуктивності. Подальші дослідження та збір додаткових даних про навантажувачі дозволять передбачити показники

продуктивності навантажувача в операції з урахуванням типу операції та типу навантажувача.

## 5. Висновки

У цьому дослідженні даних із дев'яти блоків завантаження використовувалися аналіз даних, статистичний аналіз і методи машинного навчання. Сім із цих завантажувальних одиниць надійшли з програмного забезпечення CAT Productivity, тоді як інші два – з програмного забезпечення CAT MineStar Edge. Ключові параметри виробництва були визначені за допомогою кількох змінних, які включають корисне навантаження ковша, тривалість циклу навантажувача на ковш, коефіцієнт заповнення, час циклу вантажівки, загальне корисне навантаження вантажівки та тривалість циклу навантажувача на вантажівку. Після цих аналізів було проведено аналіз виробництва до та після технічного обслуговування наземних інструментів, щоб визначити ефективність цих методів технічного обслуговування. Аналіз даних використовувався для встановлення зв'язку між середнім корисним навантаженням кожного навантажувача та тривалістю їхнього циклу. Цей зв'язок був непереконливим у визначенні оптимальної точки, коли виробництво може бути збільшено. Було виявлено велику різноманітність даних для кожного завантажувача між часом циклу та середнім корисним навантаженням. Проте три завантажувачі вказали оптимальний час циклу між 35–40 с для отримання більшого корисного навантаження.

Проведений статистичний аналіз був використаний для порівняння розподілу виробничих змінних між вантажниками, які працюють у подібних робочих умовах. Caterpillar 992K і один із завантажувальних агрегатів 990K були визнані найпродуктивнішими машинами. У наборі даних CAT MineStar Edge одна із завантажувальних одиниць Caterpillar 993K перевершила аналогічну машину 993K за всіма виробничими параметрами з невідомих причин, які стануть основою майбутніх досліджень. Вплив на ці результати може бути пов'язаний з досвідом оператора, розташуванням роботи в одній шахті та типом матеріалу.



Машинне навчання спочатку використовувалося для дослідження кореляції між змінними в наборах даних. Завдяки моделям лінійної та поліноміальної регресії, через які дані були використані, не було значної кореляції між жодною зі змінних. Хоча існує кореляція між повною плановою відстанню та часом циклу вантажівки, цей зв'язок є очікуваним, оскільки чим далі вантажівки їдуть, тим вищий час циклу. Це вказує на те, що слід провести майбутні дослідження, щоб з'ясувати, чому ця кореляція була не такою, якою вона повинна бути теоретично. З використанням чотирьох алгоритмів прогнозування алгоритм класифікатора дерева рішень дав найкращі результати в оцінці продуктивності завантажувача на основі повного аналізу всіх даних з точністю 87,99%. Наступною найкращою моделлю був випадковий ліс із точністю 81,49%. У майбутніх операціях, враховуючи тип навантажувача, ми зможемо передбачити, якими мають бути його виробничі показники.

Після генерації порівняння середнього виробництва до і після технічного обслуговування та відсоткової зміни вартості виробництва для інструментів, що займають ґрунт, на ковші навантажувача одного з 988К, було виявлено, що не було суттєвих змін у середньому виробництві шахти. Цей аналіз все ще не дає відповіді на питання, чи замінюють GET занадто часто. Майбутній аналіз спробує створити прогнозну модель для оптимальних інтервалів технічного обслуговування для GET.

Дані про продуктивність CAT значно відрізнялися від CAT MineStar Edge, оскільки вони виявилися менш детальними та значною мірою залежали від помилок людини, оскільки багато змінних потрібно було видалити. Завдання завантаження повторюється, оскільки цикл короткий (менше хвилини кожен) і виконується протягом довгих змін (від восьми до десяти годин). Надійність даних низька через введення оператором. Майбутня робота може допомогти вдосконалити моделі, описані в статті. Додатковий рік досліджень надасть більше даних і дасть більш точні моделі. Додаткова оцифровка машинної інформації заповнить програмне забезпечення

додатковими змінними та більшою різноманітністю типів машин для подальшого вдосконалення прогнозних моделей. Це також включатиме дані з одиниць кар'єрних вантажівок. Оскільки популярність дистанційного моніторингу зростає, більше компаній платитимуть за підписку на це програмне забезпечення, що дозволить використовувати більше машин для аналізу. Відвідування місця допоможе в майбутніх дослідженнях краще зрозуміти роботу та повсякденні завдання, які виконує кожен вантажник. Невідповідності в даних можна виявити не лише за допомогою цих візитів на об'єкти, але й за допомогою часових досліджень, щоб визначити перешкоди в роботі. Необхідно провести додаткові дослідження щодо кращих способів визначення того, чи GET зношені, замість того, щоб лише думка оператора вирішувала, коли їх слід змінити, оскільки це може усунути ймовірність людської помилки. Невідповідності в даних можна виявити не лише за допомогою цих візитів на об'єкти, але й за допомогою часових досліджень, щоб визначити перешкоди в роботі. Необхідно провести додаткові дослідження щодо кращих способів визначення того, чи GET зношені, замість того, щоб лише думка оператора вирішувала, коли їх слід змінити, оскільки це може усунути ймовірність людської помилки. Невідповідності в даних можна виявити не лише за допомогою цих візитів на об'єкти, але й за допомогою часових досліджень, щоб визначити перешкоди в роботі. Необхідно провести додаткові дослідження щодо кращих способів визначення того, чи GET зношені, замість того, щоб лише думка оператора вирішувала, коли їх слід змінити, оскільки це може усунути ймовірність людської помилки.

## Список літератури

1. Soofastaei, A. *Аналітика даних, застосована до гірничодобувної промисловості* ; CRC Press: Бока-Ратон, Флорида, США, 2020. [ ]
2. Молаї, Ф.; Рахімі, Е.; Сявоші, Х.; Афруз, С.Г.; Теноріо, В. Комплексний огляд Інтернету речей (IoT) та його наслідків у гірничодобувній промисловості. *Am. J. Eng. appl. Sci.* 2020 , 13 , 499–515. [ ] [ ]
3. Азіз, А.; Шелен, О.; Бодін, У. Дослідження промислового Інтернету речей для гірничодобувної промисловості: синтезована архітектура та відкриті напрямки досліджень. *IoT 2020* , 1 , 529–550. [ ] [ ]
4. Холмлунд, М.; Ван Веренберг, Ю.; Сіучита, Р.; Равальд, А.; Сарантопулос, П.; Орден Ф.В.; Закі, М. Управління клієнтським досвідом в епоху аналітики великих даних: стратегічна основа. *Ж. Автобус. рез.* 2020 , 116 , 356–365. [ ] [ ]
5. Qi, SS Управління великими даними в гірничодобувній промисловості. *Міжн. Я. Гірник. метал. Матер.* 2020 , 27 , 131–139. [ ] [ ]
6. Fekete, JA *Big Data in Mining Operations* ; Копенгагенська школа бізнесу: Фредеріксберг, Данія, 2015. [ ]
7. Хуссейн, СМ; Полрадж, MS; Нужат, С. *Зменшення джерел і мінімізація відходів* ; Elsevier: Амстердам, Нідерланди, 2021. [ ]
8. Маєр-Шенбергер, В.; Кукієр, К. *Великі дані: революція, яка змінить наше життя, роботу та мислення* ; Houghton Mifflin Harcourt: Бостон, Массачусетс, США, 2013. [ ]
9. Паньков, К.Л.; Мур, Дж. Р.; Хейл, Дж.М.; Копер К.Д.; Кубацький, Т.; Відден, К.М.; McCarter, МК Масивний зсув ґрунту на мідній шахті штату Юта генерує велику кількість геофізичних даних. *Gsa Today* 2014 , 24 , 4–9.
10. Matsimbe, J. Оптимізація продуктивності екскаватора в кар'єрах. *Міжн. J. Res. Адвент Технол.* 2020 , 8 , 1–9. [ ]

11. Nday, INM; Томас, Х. Оптимізація часу циклу для підвищення продуктивності в Ruashi Mining. *Ж. Південний. афр. ін-т Хв. метал.* 2019 , 119 , 631–638. [ ] [ ]
12. Саматемба, Б.; Чжан, Л.; Беса, Б. Оцінка та оптимізація ефективності гірничого обладнання; випадок підземної шахти Chibuluma South. *Ж. Чистий. Виробник* 2020 , 252 , 119697. [ ] [ ]
13. Паоланті, М.; Ромео, Л.; Фелічетті, А.; Манчіні, А.; Фронтоні, Е.; Лончарські, Дж. Підхід до машинного навчання для прогностичного обслуговування в промисловості 4.0. У матеріалах 14-ї міжнародної конференції IEEE/ASME з мехатронних і вбудованих систем і застосувань (MESA) 2018 р., Оулу, Фінляндія, 2–4 липня 2018 р.; С. 1–6. [ ]
14. Робатто Сімард, С.; Гамаш, М.; Дойон-Пулен, П. Сучасна практика профілактичного технічного обслуговування та очікування щодо прогнозного технічного обслуговування на шахтах Східної Канади. *Гірнична справа* 2023 , 3 , 26–53. [ ] [ ]
15. Басрі, Е.І.; Разак, ІНА; Аб-Самат, Х.; Kamaruddin, S. Планування профілактичного обслуговування (PM): огляд. *J. Qual. Технічне обслуговування інж.* 2017 , 23 , 114–143. [ ] [ ]
16. Нобахар, П.; Пурахімян Ю.; Mollaei Koshki, F. Optimum Fleet Selection Using Machine Learning Algorithms—Case Study: Zenouz Kaolin Mine. *Гірнична справа* 2022 , 2 , 528–541. [ ] [ ]
17. Бек, Дж.; Чой, Ю. Глибока нейронна мережа для прогнозування видобутку руди системами автотранспорту у відкритих шахтах. *анл. Sci.* 2020 , 10 , 1657. [ ] [ ] [ Зелена версія ]
18. Свонсон, Д. А. Про зв'язок між значеннями однієї сумарної міри похибки при використанні кількох характеристик в один і той самий момент часу: дослідження MALPE і MAPE. *екон. Фінанс.* 2015 , 5 , 1–14. [ ]
19. Caterpillar Inc. *Cat Product Link Довідковий посібник із запчастин і обслуговування* ; Caterpillar Inc.: Ірвінг, Техас, США, 2023 р. [ ]

20. Caterpillar Inc. *Посібник із вибору додатків підключення—Радіосистеми* ; Caterpillar Inc.: Ірвінг, Техас, США, 2023 р. [ ]
21. Caterpillar Inc. *Cat Product Link Technology* ; Caterpillar Inc.: Ірвінг, Техас, США, 2019. [ ]
22. MineStar Edge. Доступно в Інтернеті: <https://www.westrac.com.au/technology/minestar/minestar-fleet/minestar-edge> .
23. Продуктивність машин Катерпіллар. Доступно в Інтернеті: <https://www.westrac.com.au/technology/cat-technology/cat-productivity>.
24. Оздоган, М.; Özdoğan, Н. Сегменти часу циклу та криві розподілу часу циклу колісних навантажувачів розміру гірничодобувної промисловості — практичне дослідження. *Sci. Хв. J.* 2017 , 56 , 13–21. [ ]
25. Пісня, С.; Маркс, Е.; Pradhananga, N. Змінні впливу часу циклу самоскида для важких будівельних проектів земляних робіт. *J. Constr. інж. проект кер.* 2017 , 7 , 11–18. [ ]
26. Caterpillar Inc. *Керівництво з експлуатації та технічного обслуговування CAT: CAT MineStar Edge Production Recording* ; Caterpillar Inc.: Ірвінг, Техас, США, 2020. [ ]
27. George, B.; Nojabaei, B. Data Analyses of Quarry Operations and Maintenance Schedules: A Production Optimization Study. *Mining* 2023, 3, 347–366. <https://doi.org/10.3390/mining3020021>