

СТАТИСТИЧНИЙ АНАЛІЗ ТЕЛЕМЕТРИЧНИХ ДАНИХ МАНЕВРОВИХ ТЕПЛОВОЗІВ

Іващев Д. В.¹ [ORCID], Герасимов В. В.² [ORCID]

¹Дніпровський національний університет імені Олеся Гончара, аспірант, Україна

²Дніпровський національний університет імені Олеся Гончара,

к.т.н., доцент, Україна

Анотація. Дослідження присвячено статистичному аналізу великих масивів телеметричних даних маневрових тепловозів (понад 4,89 млн вимірювань) для контролю їхньої енергоефективності. Проаналізовано дискретні багатовимірні часові ряди, що включають рівень палива, швидкість та потужність. Встановлено, що безперервний потік даних містить апаратні та експлуатаційні перешкоди (виявлено 1501 апаратний викид), що ускладнює пряме застосування стандартних алгоритмів. Розраховано глобальні статистичні характеристики, які свідчать про високу дисперсію та асиметрію розподілу. Виявлено значну частку часу знаходження тепловозів у стані простою (57,7–98,8%), під час якого дизель-генераторна установка продовжує споживати паливо. Крос-кореляційний аналіз показав низьку загальну кореляцію між рівнем палива та потужністю (0,055–0,341), яка зростає до 0,485 у режимах активного навантаження. Доведено необхідність обов'язкової попередньої фільтрації та сегментації даних для коректного алгоритмічного виявлення експлуатаційних аномалій.

Ключові слова: телеметричні дані, маневровий тепловоз, статистичний аналіз, часовий ряд, енергоефективність, крос-кореляція, фільтрація даних.

Вступ

В умовах експлуатації маневрових тепловозів системи телеметричного моніторингу використовуються для контролю енергоефективності рухомого складу. Безперервний потік бортових даних містить аномалії, спричинені різними чинниками. Базовий (медіанний) крок дискретизації становить 1 хвилину. Вплив зовнішніх факторів та специфіка роботи ємнісних датчиків рівня палива [1] зумовлюють появу дефектів та розривів у часовому ряді. Це впливає на рівномірність масиву даних, що обмежує застосування стандартних алгоритмів визначення експлуатаційних витрат без попереднього статистичного аналізу всього часового ряду в цілому. Метою дослідження є

статистичний аналіз телеметричних даних без їх попередньої сегментації на окремі експлуатаційні стани. Це дозволяє оцінити імовірно-статистичні характеристики потоку та визначити природу аномалій у даних.

Основний матеріал

Загальний телеметричний потік формалізується у вигляді дискретного багатозначного часового ряду:

$$Z = \{t_i, V_i, S_i, P_i\}_{i=1}^N$$

де t_i — часова мітка i -го вимірювання; V_i — рівень палива у баках (л); S_i — швидкість тепловоза (км/год); P_i — потужність дизель-генераторної установки (кВт); N — кількість точок вимірювання.

Суцільна вибірка з шести маневрових тепловозів охоплює $N = 4\,898\,881$ вимірювань. Для отримання достовірних статистичних оцінок, перед розрахунком ключових моментів розподілу масив даних був очищений від аномалій. Для ідентифікації аномалій у ряді $V(t)$ застосовувалося обчислення дискретного градієнта швидкості зміни об'єму [3]:

$$\Delta V_i = \frac{V_i - V_{i-1}}{t_i - t_{i-1}}$$

Точки, у яких градієнт ΔV_i перевищує фізично можливі межі швидкості зміни об'єму палива, ідентифікувалися як апаратні викиди. У процесі обробки масиву було вилучено 1501 такий стрибок сигналу. Розраховані на очищеному масиві ключові статистичні характеристики наведені у табл. 1.

Таблица 1

Глобальні статистичні характеристики часових рядів

Характеристика (параметр)	Діапазон значень
Загальний обсяг вибірки, вимірювань	4 898 881
Медіанний рівень палива у баках V_{med} , л	3391 ÷ 4679
Частка часу знаходження у стані простою ($S = 0$), %	57,74 ÷ 98,83
Середньоквадратичне відхилення рівня палива σ_V , л	689 ÷ 1180
Максимальна зафіксована потужність ДГУ P_{max} , кВт	753 ÷ 1537
Коефіцієнт асиметрії ряду палива	-0,511 ÷ -0,246
Екссес ряду рівня палива	2,14 ÷ 4,35

Статистичний аналіз масиву швидкості $S(t)$ демонструє асиметричний розподіл. Відсоток нульових значень швидкості (стан простою) варіюється від

57,7% до 98,8% залежно від тепловоза. Водночас часовий ряд потужності $P(t)$ показує, що дизель-генераторна установка генерує енергію для допоміжних систем навіть у стані простою, демонструючи **середнє значення потужності на рівні 8,6 кВт (від 4,7 до 11,1 кВт залежно від локомотива)** із середньоквадратичним відхиленням 115–346 кВт.

Часовий ряд рівня палива $V(t)$ містить коливання, пов'язані з динамікою рідини. Розрахований коефіцієнт ексцесу досягає 4,35, що вказує на наявність значної кількості викидів розподілу [2]. Коефіцієнт асиметрії є від'ємним (від -0,24 до -0,51), що відповідає загальному тренду на споживання палива. Взаємозв'язок параметрів у часовому ряді оцінюється через крос-кореляційну функцію R_{VP} між споживанням палива та потужністю:

$$R_{VP} = \frac{\sum(V_i - \bar{V})(P_i - \bar{P})}{\sqrt{\sum(V_i - \bar{V})^2 \sum(P_i - \bar{P})^2}}$$

Розрахунки (табл. 2) показують, що глобальна крос-кореляція є низькою (0,055 – 0,341) внаслідок значної частки режимів холостого ходу.

Таблиця 2

Крос-кореляційний аналіз глобальних параметрів

Об'єкт	R (Паливо-Швидкість) загальне	R (Паливо-Потужність) загальне	R (Паливо-Потужність) в активному режимі
Тепловоз 1	0,003	0,071	0,097
Тепловоз 2	-0,011	0,114	0,258
Тепловоз 3	-0,029	0,341	0,485
Тепловоз 4	0,013	0,113	0,162
Тепловоз 5	0,003	0,151	0,135
Тепловоз 6	0,001	0,055	0,156

При виключенні ділянок простою коефіцієнт кореляції зростає до 0,485, що вказує на наявність помірного лінійного зв'язку між параметрами під навантаженням.

Висновки

Проведений статистичний аналіз масиву обсягом 4,89 млн вимірювань показує, що телеметричні ряди маневрових тепловозів характеризуються високим рівнем дисперсії (до $1,39 \times 10^6$) та наявністю асиметрії у показниках

споживання палива. Значна частка часу простою (від 57,7% до 98,8%) призводить до низької загальної крос-кореляції між рівнем палива та потужністю (0,055–0,341), яка зростає до 0,485 лише у режимах активного навантаження. Виявлення 1501 апаратного викиду та розривів зв'язку тривалістю до 176 годин підтверджує необхідність попередньої фільтрації та сегментації масиву. Виконання цих процедур є обов'язковою умовою для алгоритмічного виявлення експлуатаційних аномалій.

ЛІТЕРАТУРА

1. Navixy. Fuel level sensors: resistive, capacitive, ultrasonic methods. URL: <https://www.navixy.com/blog/fuel-level-sensors/> (дата звернення: 20.03.2026).
2. Box, G. E. P., Jenkins, G. M., Reinsel, G. C. Time Series Analysis: Forecasting and Control. 5th ed. Hoboken, NJ: John Wiley & Sons, 2015. 712 p.
3. Іващев Д., Герасимов В. Алгоритм сегментації даних вимірювань рівня палива на транспорті. Вісник Хмельницького національного університету. Технічні науки. 2025. № 2 (332). С. 60–65. DOI: 10.31891/2219-9365-2025-82-60.

STATISTICAL ANALYSIS OF TELEMETRY DATA OF SHUNTING LOCOMOTIVES

D.V. Ivashchev, V.V. Herasymov

Abstract. *The study is devoted to the statistical analysis of large arrays of telemetry data of shunting locomotives (over 4.89 million measurements) to monitor their energy efficiency. Discrete multidimensional time series including fuel level, speed, and power were analyzed. It has been established that the continuous data stream contains hardware and operational interference (1501 hardware spikes were detected), which complicates the direct application of standard algorithms. Global statistical characteristics have been calculated, indicating high variance and asymmetry of the distribution. A significant proportion of the time shunting locomotives spend in idle state (57.7–98.8%) was revealed, during which the diesel generator set continues to consume fuel. Cross-correlation analysis showed a low overall correlation between fuel level and power (0.055–0.341), which increases to 0.485 in active load modes. The necessity of mandatory preliminary filtering and segmentation of data for the correct algorithmic detection of operational anomalies is proved.*

Keywords: *telemetry data, shunting locomotive, statistical analysis, time series, energy efficiency, cross-correlation, data filtering.*

REFERENCES

1. Navixy. Fuel level sensors: resistive, capacitive, ultrasonic methods. URL: <https://www.navixy.com/blog/fuel-level-sensors/> (date of access: 20.03.2026).
2. Box, G. E. P., Jenkins, G. M., Reinsel, G. C. Time Series Analysis: Forecasting and Control. 5th ed. Hoboken, NJ: John Wiley & Sons, 2015. 712 p.
3. Ivashchev D., Herasymov V. Alhorytm sehmentatsii danykh vymiriuvan rivnia palyva na transporti. Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky. 2025. № 2 (332). S. 60–65. DOI: 10.31891/2219-9365-2025-82-60.