УДК 005.8:004.415.2 DOI: 10.31673/2412-4338.2026.029113

Ясінецький Олег Олегович
аспірант
Національний технічний університет «Харківський національний університет радіоелектроніки», Харків, Україна
ORCID 0009-0001-8771-4465
oleh.yasinetskyi@nure.ua

Фесенко Тетяна Григорівна
доктор технічних наук, професор
Національний технічний університет «Харківський національний університет радіоелектроніки», Харків, Україна
ORCID 0000-0001-9636-9598
t.fesenko@nure.ua

ЕМПІРИЧНЕ ДОСЛІДЖЕННЯ ВАРІАТИВНОСТІ ТРИВАЛОСТІ SCRUM-СПРИНТІВ ТА ЇЇ ЗВ'ЯЗКУ З ЧАСТОТОЮ ЗАТРИМОК У ВІДКРИТИХ AGILE-ПРОЄКТАХ (НА ДАНИХ TAWOS)

Анотація. *У дослідженнях ризиків затримок Scrum-проєктів аналіз будується переважно навколо обсягових метрик - velocity, story points та їхніх похідних. Календарний ритм ітерацій як самостійний емпіричний предмет опрацьований слабше. Стаття зміщує фокус на фактичну календарну тривалість Scrum-спринтів та її проєктну варіативність. Емпіричну основу склали 4335 закритих спринтів з 36 open-source проєктів TAWOS за період 2012–2020 рр., відібраних після чотириетапної процедури фільтрації. Для кожного спринту порівняно заплановану та фактичну тривалість; на рівні проєкту обчислено коефіцієнт варіації фактичної тривалості (CV) як міра дисципліни календарного ритму. У вибірці 58,02% спринтів завершено в межах припустимого відхилення (±10%), 22,58% - із суттєвою затримкою (понад 25%), 14,33% - із незначною затримкою (10–25%), 5,07% - із випередженням плану. На підвибірці з 31 проєкту, у яких налічувалося не менш як 30 спринтів, виявлено статистично значущий ранговий зв'язок CV фактичної тривалості з часткою спринтів зі суттєвою затримкою (Spearman $\rho = 0,609$, $p = 0,000275$), а також зворотний зв'язок з часткою вчасних спринтів ($\rho = –0,538$, $p = 0,002$). Розподіли фактичних тривалостей між проєктами виявилися виражено неоднорідними (Kruskal–Wallis $H = 1048,70$, $p < 0,001$). Тертильна стратифікація за CV дала три групи проєктів - «дисципліновані», «помірні», «хаотичні», - які статистично розрізняються за часткою затримок ($H = 8,03$, $p = 0,018$). За період спостережень частка спринтів зі суттєвою затримкою монотонно знижується ($\rho = –0,767$, $p = 0,016$), що може свідчити про процес дозрівання Agile-практик у спільноті відкритих проєктів. Отримані результати позиціонуються як емпіричний baseline для подальших предиктивних моделей у дослідженнях ризиків затримок Scrum-проєктів, не як самостійний прогностичний інструмент. Запропонована тришарова типологія може використовуватися як описова стратифікаційна рамка для первинної оцінки календарної дисципліни проєктів IT-аутсорсингу.*

Ключові слова: *Scrum, варіативність тривалості, TAWOS, agile-проєкти, статистичний аналіз, коефіцієнт варіації, типологія проєктів, ризики затримок.*

Oleh Yasinetskyi
PhD student
National Technical University "Kharkiv National University of Radio Electronics", Kharkiv, Ukraine
ORCID 0009-0001-8771-4465
oleh.yasinetskyi@nure.ua

Tetiana Fesenko
Doctor of Technical Sciences, Professor
National Technical University "Kharkiv National University of Radio Electronics", Kharkiv, Ukraine
ORCID 0000-0001-9636-9598
t.fesenko@nure.ua

AN EMPIRICAL STUDY OF SCRUM SPRINT DURATION VARIABILITY AND ITS RELATIONSHIP WITH DELAY FREQUENCY IN OPEN-SOURCE AGILE PROJECTS (USING THE TAWOS DATASET)

Abstract. *Research on delay risk in Agile projects tends to emphasise story points, velocity, and issue-level signals, with calendar duration of sprints rarely treated as a project-level variable in its own right. The present study addresses this gap. Using the open TAWOS dataset, we examine the actual calendar duration of Scrum sprints, its variability across projects, and the link between that variability and how often projects miss planned dates. After a four-step filtering procedure the working sample contains 4335 closed sprints from 36 open-source projects over the 2012–2020 period. For every sprint, planned and actual duration are compared. At the project level, the coefficient of variation of actual duration (CV) serves as a measure of calendar rhythm discipline. Within the sample, 58.02% of sprints fall within ±10% of the planned duration; 22.58% exceed the plan by more than 25%; 14.33% show minor delays of 10–25%; and 5.07% finish ahead of schedule. Across the 31 projects with at least 30 sprints each, CV of actual duration is*

© 2026 Ясінецький О.О., Фесенко Т.Г. Цей матеріал ліцензовано за умовами CC BY 4.0. https://creativecommons.org/licenses/by/4.0/

positively associated with the share of major-delay sprints (Spearman ρ = 0.609, p = 0.000275) and inversely with the share of on-time sprints (ρ = –0.538, p = 0.002). The Kruskal–Wallis test indicates that sprint duration distributions differ substantially across projects (H = 1048.70, p < 0.001). Tertile stratification by CV produces three groups - disciplined, moderate, chaotic - that also differ by delay share (H = 8.03, p = 0.018). The share of major-delay sprints declines monotonically over the observation window (ρ = –0.767, p = 0.016), which may indicate the gradual maturation of Agile practices in the open-source community. The findings are positioned as an empirical baseline for future predictive work on Scrum delay risk, not as a stand-alone prognostic model. The proposed three-tier typology may be used as a descriptive stratification frame for initial assessment of calendar discipline in IT-outsourcing projects.

Keywords: *Scrum, duration variability, TAWOS, agile projects, statistical analysis, coefficient of variation, project typology, delay risks.*

1. Вступ

У Scrum-проєктах планування і постачання цінності організовано довкола ітерацій фіксованої тривалості - спринтів. На скільки добре команда тримається запланованого ритму, видно одразу на двох вимірах: за виконаним обсягом і за календарними термінами. Дослідження ризиків затримок зосереджені переважно на першому вимірі - обсягових метриках на кшталт velocity і story points [1, 6]. Календарна тривалість спринту як предмет емпіричного аналізу опрацьована слабше, попри те, що проєктні менеджери на практиці орієнтуються радше на дотримання дат, ніж на обсяги.

Попередні роботи автора [2, 3] стосувалися інтегральної оцінки ризиків затримок Scrum-проєктів через композитний індекс на основі метрик продуктивності та цифрової комунікації. У тих роботах варіативність тривалості спринтів окремо не виокремлювалася й латентно враховувалася через індекси обсягових змінних. Натомість досвід IT-аутсорсингу свідчить, що нестабільність календарного ритму є самостійним сигналом ризику: команди з нестабільним календарним ритмом ітерацій частіше демонструють непрогнозованість постачання, навіть коли формальні обсягові нормативи виконуються.

У цьому дослідженні емпірично характеризується варіативність тривалості Scrum-спринтів та її статистичний зв'язок з частотою затримок на масштабному масиві відкритого датасету. У межах цієї статті предиктивна модель не розробляється. Робота формує описовий емпіричний фундамент, на якому подальші predictive-моделі (зокрема, представлена у [3]) можуть бути обґрунтовано побудовані та інтерпретовані.

2. Постановка проблеми

Незважаючи на широке практичне застосування Scrum у проєктах розробки програмного забезпечення, у літературі з прогнозування ризиків затримок переважає аналіз обсягових метрик - velocity, story points та їхніх похідних [1, 6]. Натомість сама календарна тривалість спринтової ітерації, попри її пряму прив'язку до дат постачання й релізного планування, як предмет окремого емпіричного аналізу опрацьована значно слабше. Менеджерські рішення в IT-аутсорсингу значною мірою спираються на дотримання дат: пропущений дедлайн релізу впливає на контрактні зобов'язання, маркетингові плани замовника та продуктові roadmap-и. Через це характеристика варіативності календарного ритму спринтів і встановлення її зв'язку з частотою серйозних затримок становить актуальну наукову задачу для теорії та практики управління Scrum-проєктами. Проблема полягає у відсутності систематичної емпіричної характеризації цього зв'язку на репрезентативних масштабних даних.

3. Аналіз останніх досліджень і публікацій

Накопичення відкритих репозиторіїв з даними про життєвий цикл Agile-задач створило передумови для емпіричних досліджень масштабу, недосяжного у попередні десятиліття. Tawosi V. зі співавторами [1] представили відкритий датасет TAWOS, що містить дані про понад 500 тис. задач та десятки тисяч спринтів з відкритих agile-проєктів. Структура датасету включає три ключові таблиці - Issue, Sprint та Change_Log. Це відкриває можливість аналізу патернів планування на рівні окремої ітерації. У дослідженнях, побудованих на TAWOS, превалюють задачі прогнозування story points [4] та виявлення дефектів, тоді як тривалість спринтової ітерації як окремий емпіричний предмет залишається у фоновій ролі.

У дослідженнях Kula E. та колег [6, 7] on-time delivery постає центральним об'єктом аналізу. Автори детально показують, як організаційні та контекстуальні чинники впливають на терміни постачання фіч у промислових Agile-командах. Проте сама календарна тривалість спринту в цих роботах не виокремлена як вимірюваний предиктор. Робота Choetkiertikul M. et al. [5] про прогнозування затримок у програмних проєктах через networked classification оперує рівнем окремих issue. Агрегування до рівня спринту як одиниці аналізу там відсутнє. Спільне для цих робіт - фокус на затримках як виході моделі, а не на самій структурі ітераційного ритму як змінній, що цей вихід пояснює.

Емпіричні дослідження психо-соціальних аспектів Scrum, що використовують Mining Software Repositories як методологію (Ortu M. et al. [8]), концентруються на текстовому аналізі коментарів та emotion mining. Календарний ритм там не аналізується. Систематичний огляд метрик в Agile/Lean розробці (Kupiainen E., Mäntylä M.V., Itkonen J. [13]) узагальнює значну кількість досліджень метрик продуктивності та якості - однак і в цьому огляді календарна тривалість спринту як самостійний предмет емпіричного аналізу майже не фігурує. Дослідження velocity variability у стабільних командах (огляди методології Scrum [9, 10]) фіксують природну варіативність обсягових метрик у межах 15–30%, проте знов оминають аналіз тривалостей як окремий феномен.

З проведеного аналізу випливають три прогалини у поточному стані досліджень. Літературний фокус усе ще зосереджений на обсягових метриках, тоді як практика управління постачанням значно більше орієнтована на дотримання дат. Агрегування результатів на рівні проєкту (а не окремої задачі чи спринту) залишається порівняно рідкісним: дисципліна календарного ритму є передусім проєктною характеристикою, не задачною. І нарешті, TAWOS - попри широке цитування в дослідженнях story points та defect prediction - ще не використовувався для систематичної статистичної характеризації тривалостей спринтів та її зв'язку з частотою затримок. Ці три прогалини й окреслюють предметне поле поточного дослідження.

4. Мета і задачі дослідження

Метою дослідження є емпірична статистична характеристика варіативності календарної тривалості Scrum-спринтів та виявлення її статистичного зв'язку з частотою затримок ітерацій на масштабному масиві даних відкритих agile-проєктів.

Для досягнення мети сформульовано такі завдання:

– визначити дескриптивну статистичну характеристику запланованої та фактичної календарної тривалості Scrum-спринтів у вибірці датасету TAWOS;

– побудувати класифікацію спринтів за рівнем відхилення фактичної тривалості від запланованої та оцінити частку кожної категорії;

– обчислити коефіцієнт варіації фактичної тривалості на рівні окремого проєкту та статистично оцінити його зв'язок з частотою затримок;

– розробити емпіричну типологію проєктів за рівнем дисципліни календарного ритму та статистично верифікувати її розрізняльну здатність;

– охарактеризувати темпоральні тренди частоти затримок у вибірці за період 2012–2020 рр.

5. Матеріали і методи дослідження

5.1. Джерело даних і процедура формування вибірки

Емпіричною основою дослідження виступив відкритий датасет TAWOS [1] - він містить дані про життєвий цикл задач та спринтів з відкритих agile-проєктів. Для аналізу використано таблицю Sprint з її полями стану ітерації (state), запланованих дат початку і завершення, фактичної дати завершення. Стан спринту визначається у термінах Jira: CLOSED - завершений, ACTIVE - поточний, FUTURE - запланований. У дослідженні аналізувалися виключно спринти зі станом CLOSED - лише для них наявні всі три ключові часові показники.

Фінальна робоча вибірка формувалася чотирма послідовними кроками. Спочатку з вихідного набору відбиралися записи зі станом CLOSED. Далі виключалися записи з пропущеними значеннями полів початку, запланованого завершення та фактичного завершення. Третій крок - валідація календарних дат: виключалися записи з некоректним роком (наприклад, артефакти синтаксичного парсингу типу року 0015) та записи з датою фактичного завершення 1970-01-01 (Unix epoch як індикатор серіалізованого NULL). Четвертий крок - фільтрація за відомими операційними нормами Scrum: запланована тривалість обмежена діапазоном 3–60 календарних днів (це відповідає типовому діапазону ітерацій від одного тижня до двох місяців), фактична - 0–120 днів (для виключення артефактів адміністративного закриття прострочених спринтів через значний час після фактичного завершення робіт). Sensitivity-перевірка показала, що виключення записів з actual_days < 1 не змінює напряму та статистичної значущості основних встановлених зв'язків. Критерії фільтрації та обсяги вибірки на кожному кроці узагальнено у таблиці 1.

Таблиця 1

Процедура формування вибірки

Крок / підсумок	Критерій	Залишилося спринтів
Початковий обсяг	Записи Sprint, успішно зчитані після технічного імпорту CSV	4537
Крок 1: стан CLOSED	state = CLOSED	4426
Крок 2: повнота дат	start_date, end_date, complete_date не є NULL	4426
Крок 3: валідність дат	Виключення некоректних років і артефактів Unix epoch	4422
Крок 4: фільтр тривалості	`3 ≤ planned_days ≤ 60; 0 ≤ actual_days ≤ 120`	4335
Підсумок	Спринти, придатні для аналізу	4335

Окремий методологічний нюанс: поле complete_date у датасеті TAWOS відображає момент адміністративного закриття спринту в системі Jira, а не завжди реальне завершення робіт командою. Цей операційний нюанс розглянуто як обмеження дослідження у розділі 6. Уточнення щодо термінології: у класичній інтерпретації Scrum Guide спринт є подією фіксованої тривалості (fixed-length event); у даному дослідженні

аналізується фактична тривалість, зафіксована у системі трекінгу задач, що може відображати як процесну дисципліну команди, так і особливості адміністрування ітерацій.

Фінальна робоча вибірка - 4335 спринтів з 36 проєктів за період з 28 лютого 2012 р. по 12 жовтня 2020 р. (9 календарних років). Для аналізу проєктних патернів додатково сформовано підвибірку з проєктів, у яких налічувалося не менш як 30 спринтів (для стабільності оцінок коефіцієнта варіації); вона складається з 31 проєкту та 4273 спринтів.

5.2. Операціоналізація змінних

На рівні окремого спринту і обчислено чотири похідні величини: заплановану тривалість P_i, фактичну тривалість A_i, абсолютне D_i та відносне Δ_i відхилення. Розрахункові співвідношення:

$$P_i = E_i - S_i \quad (1)$$

$$A_i = C_i - S_i \quad (2)$$

$$D_i = A_i - P_i \quad (3)$$

$$\Delta_i = \frac{A_i - P_i}{P_i} \times 100\% \quad (4)$$

де S_i - запланована дата початку ітерації; E_i - запланована дата завершення ітерації; C_i - фактична дата завершення ітерації, зафіксована у системі трекінгу задач Jira (поле complete_date); P_i - запланована календарна тривалість; A_i - фактична календарна тривалість; D_i - абсолютне календарне відхилення; Δ_i - відносне календарне відхилення, %. Усі тривалості виражено у календарних днях.

На основі Δ_i побудовано порядкову класифікацію виходу спринту за чотирма категоріями: early ($\Delta_i < -10\%$), on-time ($-10\% \le \Delta_i \le 10\%$), minor-delay ($10\% < \Delta_i \le 25\%$), major-delay ($\Delta_i > 25\%$). Граничні значення ±10% та 25% - операційні евристики припустимого календарного відхилення, узгоджені з практикою аналізу on-time delivery [6]. Чутливість основних результатів до зміни цих порогів - окрема тема для подальших досліджень. Ці пороги не мають нормативного статусу; вони використовуються лише як аналітична рамка для впорядкування спостережень.

Для агрегування на рівні проєкту j обчислено три показники: коефіцієнт варіації фактичної тривалості CV_j як міра дисципліни календарного ритму, частку спринтів зі суттєвою затримкою M_j та частку вчасних спринтів O_j:

$$CV_j = \frac{\sigma(A_j)}{\mu(A_j)} \quad (5)$$

$$M_j = \frac{n_j^{major}}{n_j} \times 100\% \quad (6)$$

$$O_j = \frac{n_j^{on}}{n_j} \times 100\% \quad (7)$$

де $\sigma(A_j)$ - стандартне відхилення фактичної тривалості спринтів проєкту j; $\mu(A_j)$ - середнє арифметичне фактичної тривалості; n_j - загальна кількість спринтів у проєкті j; n^{major}_j, n^{on}_j - кількість спринтів проєкту j, що належать відповідно до категорій major-delay та on-time. Аналогічно обчислено частки minor-delay та early. Додатково для характеристики стабільності планувального горизонту обчислено CV запланованої тривалості за тією ж формулою (5), але із заміною A_j на P_j.

5.3. Статистичний інструментарій

Для дескриптивної характеризації використано середнє, медіану, моду, стандартне відхилення, інтерквартильний розмах, а також показники асиметрії та ексцесу як індикатори відхилення розподілу від нормального. Зв'язок коефіцієнта варіації з частотою затримок оцінено коефіцієнтом рангової кореляції Spearman (ρ): попередній огляд даних підтвердив значну позитивну асиметрію розподілу Δ_i та невиконання припущення нормальності, тому ранговий метод доречніший за параметричний. Міжпроєктну однорідність тривалостей верифіковано непараметричним H-тестом Kruskal–Wallis, який не вимагає припущення нормальності та придатний для порівняння розподілів між понад двома незалежними групами. Зауваження щодо застосування цього тесту: спринти у межах одного проєкту можуть бути автокорельованими, тому H-тест використано як непараметричний тест міжпроєктної неоднорідності, а не як повну модель ієрархічної залежності; врахування внутрішньопроєктних кореляцій - предмет окремих подальших досліджень. Стратифікація проєктів виконана методом тертильного поділу за значенням CV_j для прозорості та інтерпретованості процедури; ця процедура розглядається як описова стратифікація, а не кластерний алгоритм у машинно-навчальному сенсі. Рівень статистичної значущості - $\alpha = 0{,}05$.

6. Результати дослідження

6.1. Дескриптивна характеристика тривалостей спринтів

Дескриптивні статистики запланованої та фактичної тривалості спринтів узагальнено у таблиці 2. Мода та медіана запланованої тривалості - рівно 14 календарних днів. Це відповідає поширеній практиці двотижневих Scrum-ітерацій. Середнє арифметичне запланованої тривалості (13,64 дня) практично збігається з медіаною: команди зазвичай не планують спринти суттєво коротшими чи довшими за двотижневі.

Фактична тривалість має суттєво ширший розкид. Стандартне відхилення фактичної тривалості (11,92 дня) удвічі перевищує таке для запланованої (5,73 дня), а максимальне зафіксоване значення сягає 111 днів.

Медіанне абсолютне відхилення фактичної тривалості від запланованої - лише 0,28 дня (близько 7 годин), тоді як середнє арифметичне такого відхилення - 3,31 дня. Розбіжність прямо вказує на сильно правосторонньо асиметричний розподіл відхилень: половина спринтів завершується практично точно в плановий день, інша половина демонструє широкий хвіст затримок, який «розтягує» середнє.

Кількісне підтвердження цього характеру розподілу - показники асиметрії (skewness) = 6,30 та ексцесу (kurtosis) = 59,43 для Δ_i. Обидва значення значно перевищують пороги нормального розподілу (skewness = 0, kurtosis = 0) і вказують на необхідність непараметричних методів аналізу. Графічне відображення розподілу абсолютних відхилень фактичної тривалості від запланованої наведено на рисунку 1.

Таблиця 2

Дескриптивна статистика тривалостей спринтів

Показник	Запланована	Фактична	Відхилення, днів
N	4335	4335	4335
Середнє	13,64	16,96	3,31
Медіана	14,00	14,00	0,28
Мода (днів)	14	14	-
Стандартне відхилення	5,73	11,92	10,16
Мінімум	3,00	0,00	–28,11
Максимум	60,00	111,09	98,09
Q1 (нижній квартиль)	11,11	13,07	–0,03
Q3 (верхній квартиль)	14,00	16,73	2,85
Інтерквартильний розмах	2,89	3,66	2,88

Рис. 1. Гістограма розподілу абсолютних відхилень фактичної тривалості від запланованої (4335 спринтів). Вісь X обрізано на рівні 60 днів для покращення читабельності

6.2. Класифікація спринтів за рівнем відхилення

Розподіл спринтів за категоріями виходу подано у таблиці 3. Близько 58% спринтів вибірки завершено у припустимих межах ±10% від запланованої тривалості - найбільша й домінантна категорія. Друга за чисельністю - major-delay (22,58%): понад п’яту частину спринтів у відкритих agile-проєктах виходить за межі прийнятного відхилення на 25% або більше. Категорія minor-delay (14,33%) описує проміжний стан. Випередження плану (early) залишається відносно рідкісним явищем (5,07%).

Сума категорій з відхиленням понад 10% (minor-delay та major-delay) - 36,91%. Це й є агрегована частка спринтів, які виходять за межі прийнятного календарного відхилення. Така частка співмірна з емпіричними результатами досліджень on-time delivery у large-scale agile software development [6], хоча пряме порівняння ускладнене відмінностями в одиниці аналізу та операціональному визначенні затримки.

Таблиця 3

Розподіл спринтів за категоріями виходу

Категорія	Діапазон Δ_i	Кількість	Частка
Early	$\Delta_i < -10\%$	220	5,07%
On-time	$-10\% \le \Delta_i \le 10\%$	2515	58,02%
Minor-delay	$10\% < \Delta_i \le 25\%$	621	14,33%
Major-delay	$\Delta_i > 25\%$	979	22,58%
Разом	-	4335	100,00%

6.3. Міжпроєктна варіативність тривалості та її зв'язок з частотою затримок

Міжпроєктна структура тривалостей виявила сильно виражену гетерогенність. H-тест Kruskal–Wallis для фактичної тривалості між проєктами (k = 31, n = 4273) дав значення H = 1048,70 при p < 0,001. Це свідчить про статистично значущі відмінності між розподілами фактичних тривалостей у різних проєктах. Іншими словами, фактична тривалість спринтів є передусім характеристикою проєкту, а не однаковою характеристикою для всіх проєктів вибірки.

Розкид коефіцієнта варіації фактичної тривалості CV_j серед проєктів - від 0,119 до 0,952 (медіана 0,500). Розкид частки спринтів зі суттєвою затримкою M_j - від 4,60% до 81,63%. Розкид частки вчасних спринтів O_j - від 6,12% до 86,62%. На одному кінці спектру існують проєкти, у яких практично 9 з 10 спринтів виходять за межі прийнятного відхилення; на іншому - проєкти, де ця частка не перевищує 10%.

Ранговий кореляційний аналіз Spearman за 31 проєктом виявив сильний позитивний зв'язок між CV_j та M_j ($\rho = +0{,}609$, p = 0,000275), а також зворотний зв'язок між CV_j та O_j ($\rho = -0{,}538$, p = 0,002). Параметричний коефіцієнт Pearson для тієї ж пари змінних дає r = +0,491 (p = 0,005) - дещо слабкіше за непараметричний еквівалент, що узгоджується з очікуваною непараметричною природою зв'язку. Зв'язок між плановою та фактичною варіативностями тривалостей виявився помірним позитивним ($\rho = +0{,}511$, p = 0,003). Це означає, що частина варіативності тривалостей закладається ще на етапі планування, проте значна її частка з'являється саме на стадії виконання.

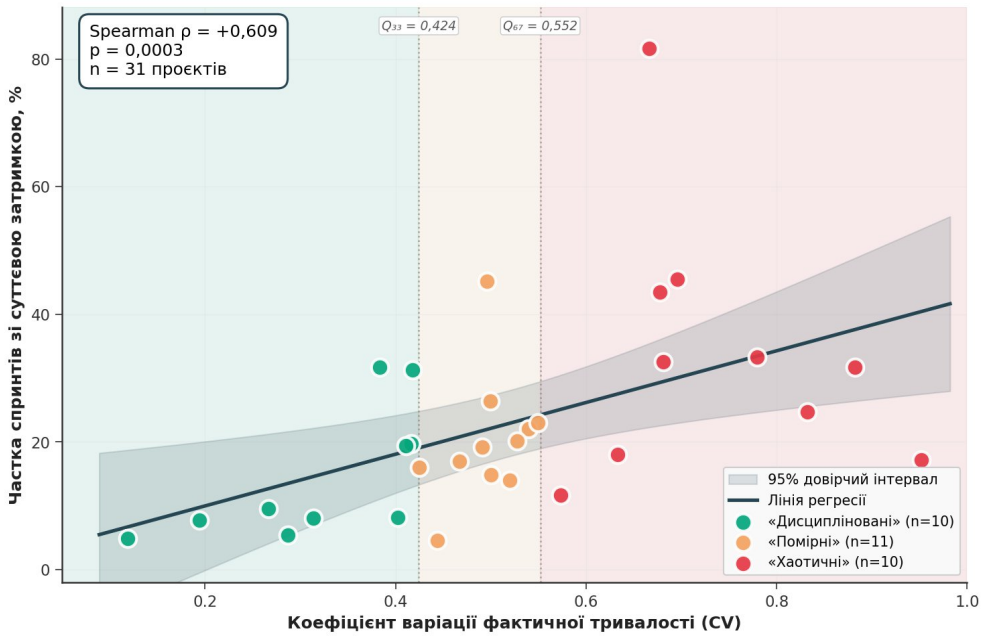

Рис. 2. Зв'язок коефіцієнта варіації фактичної тривалості з часткою спринтів зі суттєвою затримкою (n = 31 проєкт; Spearman $\rho = +0{,}609$, p < 0,001)

Графічне представлення основного зв'язку наведено на рисунку 2. Реальні дані проєктів формують виражену впорядковану структуру: проєкти з низькою фактичною варіативністю групуються в нижньому лівому квадранті (низький M_j), проєкти з високою варіативністю - у верхньому правому. Таблиця 4 узагальнює характеристики десяти найбільших за обсягом проєктів вибірки.

Таблиця 4

Топ-10 проєктів вибірки за обсягом спринтів

ID	N	Plan. med.	Act. med.	CV act.	% on-time	% major
22	542	13,0	13,85	0,832	62,55%	24,72%
21	448	11,0	13,93	0,882	52,90%	31,70%
28	356	26,4	29,15	0,550	52,53%	23,03%
12	298	14,0	14,63	0,528	41,61%	20,13%
36	296	14,0	13,96	0,500	79,39%	14,87%
4	224	14,0	14,00	0,491	73,21%	19,20%
11	206	7,0	7,05	0,573	67,48%	11,65%
13	161	14,0	14,04	0,633	56,52%	18,01%
34	143	17,1	21,10	0,696	35,66%	45,46%
3	142	14,0	13,98	0,194	86,62%	7,75%

6.4. Типологія проєктів за рівнем дисципліни календарного ритму

Тертильний поділ 31 проєкту за значенням CV_j дозволив побудувати тришарову типологію T_j як описову стратифікаційну рамку (не кластерний алгоритм у машинно-навчальному сенсі). Граничні значення CV для розмежування - 0,424 (нижня третина) та 0,552 (верхня третина). Формально типологія описується кусково-визначеною функцією належності:

$$T_j = Disciplined, якщо\ CV_j < 0{,}424\ T_j = Moderate, якщо\ 0{,}424 \le CV_j < 0{,}552\ T_j = Chaotic, якщо\ CV_j \ge 0{,}552 \quad (8)$$

Запропоновані категорії пов'язано з характером дотримання календарного ритму:

- «дисципліновані» проєкти (10 проєктів, $CV_j < 0{,}424$) - характеризуються найвищою регулярністю тривалостей; середня частка вчасних спринтів становить 67,17%, середня частка major-delay - 14,61%;
- «помірні» проєкти (11 проєктів, $0{,}424 \le CV_j < 0{,}552$) - займають проміжне положення; середня частка вчасних - 58,88%, середня частка major-delay - 20,22%;
- «хаотичні» проєкти (10 проєктів, $CV_j \ge 0{,}552$) - мають найвищу варіативність тривалостей; середня частка вчасних спринтів - лише 43,69%, середня частка major-delay - 33,97%.

Розрізняльна здатність типології перевірена методом Kruskal–Wallis для частки major-delay між трьома кластерами: H = 8,03, p = 0,018. Отже, типологія статистично розрізняє проєкти за поведінкою затримок, не є артефактом метричного поділу. Структуру кластерів узагальнено у таблиці 5.

Таблиця 5

Емпірична типологія проєктів за дисципліною календарного ритму

Тип	Діапазон CV	К-сть проєктів	Спринтів	Серед. % on-time	Серед. % major
Disciplined	< 0,424	10	804	67,17%	14,61%
Moderate	0,424–0,552	11	1679	58,88%	20,22%
Chaotic	≥ 0,552	10	1790	43,69%	33,97%
Усі	-	31	4273	-	-

6.5. Темпоральні тренди

Річна динаміка частки спринтів кожної категорії виходу показала виражений тренд: частка спринтів зі суттєвою затримкою монотонно знижується протягом періоду спостережень. У 2013 р. ця частка становила 36,1%, у 2014 - 32,4%, у 2017 - 16,2%, у 2018 - 15,3%. Ранговий тест Spearman для пари (рік, M_j) дає $\rho = -0{,}767$, $p = 0{,}016$ - статистично значуще монотонне зниження. Симетричний тренд для частки on-time спринтів ($\rho = +0{,}567$) не досягає рівня статистичної значущості при поточному $n = 9$ ($p = 0{,}112$). Це, ймовірно, пов'язано з перерозподілом між сусідніми категоріями: зниження major-delay частково компенсує зростання minor-delay, не лише on-time.

Графічне представлення темпоральних трендів подано на рисунку 3.

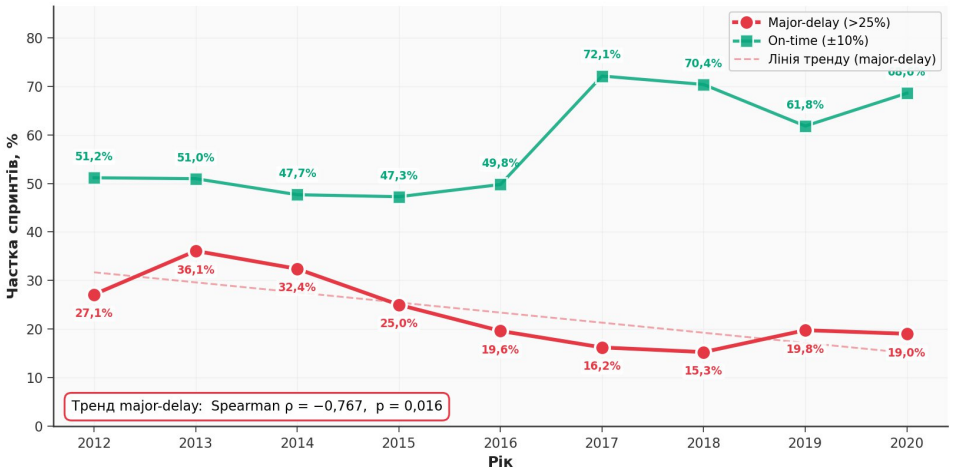

Рис. 3. Темпоральна динаміка частки спринтів зі суттєвою затримкою у вибірці TAWOS, 2012–2020 рр.

7. Обговорення результатів

Отримані результати дозволяють зробити кілька узагальнень про природу спринтового ритму у відкритих agile-проєктах. Незважаючи на широке припущення про «нормальну» варіативність ітераційної тривалості, емпіричні дані свідчать про інший характер розподілу: половина спринтів вибірки завершується практично точно у плановий день (медіана відхилення - 0,28 дня), тоді як інша половина утворює широкий правосторонньо асиметричний хвіст. Використання середніх показників відхилення стає малоінформативним, а опора на квантильні характеристики чи непараметричні методи стає необхідністю для подальших досліджень.

Виявлений сильний позитивний зв'язок коефіцієнта варіації фактичної тривалості з часткою спринтів зі суттєвою затримкою ($\rho = +0{,}609$, $p < 0{,}001$) є самостійним емпіричним результатом, що відкриває окреме предметне поле для обговорення. Дисципліна календарного ритму - не лише непрямий індикатор стабільності процесу. Це характеристика, тісно пов'язана з частотою серйозних відхилень від планів. Результат узгоджується з ширшим припущенням flow-oriented підходів про те, що зростання процесної варіативності пов'язане зі зниженням передбачуваності виконання. Емпірична перевірка цього припущення на рівні спринтового ритму у відкритих agile-проєктах становить окремий внесок дослідження.

Побудована тришарова типологія проєктів («дисципліновані» - «помірні» - «хаотичні») демонструє, що відмінності між цими групами не просто метричні: вони статистично значущо проявляються на рівні частки серйозних затримок ($H = 8{,}03$, $p = 0{,}018$). Це дає підставу розглядати запропоновану типологію як просту інтерпретовану евристику для первинної стратифікації проєктів за календарною дисципліною. Її прогностичне використання у задачах класифікації або раннього попередження потребує додаткової перевірки на незалежних даних і не входить до завдань поточної роботи.

Монотонне негативне зниження частки major-delay у часі ($\rho = -0{,}767$, $p = 0{,}016$) узгоджується з гіпотезою про дозрівання Agile-практик у спільноті open-source проєктів. У ранні роки спостереження (2012–2014) команди частіше демонстрували значні відхилення від плану, у пізніші (2017–2018) - рідше. Це може відображати поступове вдосконалення інструментарію оцінювання, поширення стандартизованих практик планування та засвоєння спільнотою накопиченого досвіду. Подальше дослідження може зосередитися на ідентифікації конкретних механізмів цього тренду - чи це ефект досвіду команд, ефект інструментальної інфраструктури, чи селекційний ефект «виживаності» більш дисциплінованих проєктів.

Обмеження дослідження пов'язані передусім з його дескриптивною природою. Виявлені кореляції є статистичними, а не каузальними: висока варіативність тривалості та висока частка затримок можуть мати спільну першопричину (наприклад, погану дисципліну планування у певних командах) без прямого причинного зв'язку між собою. Друге обмеження - характер поля complete_date у TAWOS: воно фіксує момент адміністративного закриття спринту в Jira, що не завжди збігається з реальним моментом завершення робіт. Частина зафіксованих «суттєвих затримок» може бути артефактом запізнілого адміністративного закриття, не реальною операційною затримкою. Третє - характер вибірки TAWOS як добірки відкритих agile-проєктів: підстави для узагальнення на закриті комерційні проєкти IT-аутсорсингу потребують додаткового обґрунтування у наступних роботах. Четверте - використання лише календарної тривалості без обсягових метрик (story points, completion rate), що становило свідомий методологічний вибір для виокремлення предмета аналізу.

Поточні результати інтегруються у дослідницьку програму, представлену у попередніх роботах автора [2, 3], як описовий емпіричний фундамент. Робота [2] запропонувала аналітичний огляд методологічних підходів до управління ризиками IT-проєктів. Робота [3] - інформаційну технологію інтегрованої оцінки ризиків з використанням композитного індексу на основі продуктивнісних та комунікаційних метрик. Поточна стаття не пропонує нової предиктивної моделі і не використовує комунікаційних метрик, не будує інтегрального індексу ризику, не здійснює класифікаційних прогнозів. Її внесок обмежено емпіричною характеризацією календарної

варіативності спринтів як потенційної ознаки для майбутніх моделей. У цій логіці поточне дослідження виконує роль baseline-фундаменту, не альтернативи моделям, представленим у попередніх роботах.

8. Висновки та перспективи подальших досліджень

У роботі виконано емпіричне статистичне дослідження варіативності календарної тривалості Scrum-спринтів та її зв'язку з частотою затримок на масштабному масиві відкритого датасету TAWOS. Робоча вибірка налічувала 4335 закритих спринтів з 36 open-source проєктів за період з 2012 по 2020 роки.

Основні наукові результати:

1. Емпірично встановлено, що 58,02% спринтів вибірки завершуються у припустимих межах відхилення (±10% від запланованої тривалості), 22,58% - із суттєвою затримкою (понад 25%), 14,33% - із незначною затримкою, 5,07% - із випередженням. Розподіл відхилень характеризується сильною правосторонньою асиметрією (skewness = 6,30) та значним позитивним ексцесом (kurtosis = 59,43), що вимагає застосування непараметричних методів у подальших дослідженнях.

2. Виявлено статистично значущий позитивний ранговий зв'язок коефіцієнта варіації фактичної тривалості з часткою спринтів зі суттєвою затримкою на рівні проєкту (Spearman $\rho = +0{,}609$, $p = 0{,}000275$; $n = 31$), а також зворотний зв'язок з часткою вчасних спринтів ($\rho = -0{,}538$, $p = 0{,}002$). Це підтверджує, що дисципліна календарного ритму є самостійною характеристикою проєкту, тісно пов'язаною з частотою серйозних відхилень.

3. Тестом Kruskal–Wallis встановлено вкрай виражену міжпроєктну гетерогенність фактичних тривалостей ($H = 1048{,}70$, $p < 0{,}001$) - тривалість спринту є передусім характеристикою конкретного проєкту, а не глобальною константою практики.

4. Розроблено тришарову емпіричну типологію проєктів за рівнем дисципліни календарного ритму («дисципліновані», «помірні», «хаотичні») як описову стратифікаційну рамку, розрізняльна здатність якої статистично верифікована ($H = 8{,}03$, $p = 0{,}018$). Типологія базується на операціональній ознаці (коефіцієнт варіації фактичної тривалості) та може використовуватися як проста інтерпретована евристика для первинної стратифікації проєктів; її прогностичне застосування потребує додаткової перевірки на незалежних даних.

5. Виявлено статистично значуще монотонне зниження частки спринтів зі суттєвою затримкою у вибірці за період 2012–2020 рр. ($\rho = -0{,}767$, $p = 0{,}016$), що може свідчити про процес дозрівання Agile-практик у спільноті відкритих проєктів.

Подальші дослідження доцільно зорієнтувати на каузальний аналіз механізмів виявленого зв'язку варіативності з частотою затримок, верифікацію результатів на даних закритих комерційних проєктів IT-аутсорсингу, інтеграцію розробленої типології як ознаки у предиктивних моделях управління ризиками Scrum-проєктів.

Внесок авторів

Ясінецький О.О. - концептуалізація, методологія дослідження, обробка та статистичний аналіз даних TAWOS, підготовка візуалізацій, написання основного тексту статті. Фесенко Т.Г. - наукове керівництво, концептуалізація, аналіз джерел та підготовка огляду літератури, критичне рецензування та редагування тексту.

Декларація про штучний інтелект

При підготовці статті використовувалися інструменти штучного інтелекту (Claude, ChatGPT) для допомоги у редагуванні стилістики україно- та англомовного тексту, генерації допоміжного коду на Python для статистичної обробки даних TAWOS, оформлення таблиць і списку літератури. Усі змістовні наукові твердження, постановка задачі, інтерпретація отриманих результатів та висновки належать авторам. Емпіричні дані, статистичні розрахунки (Spearman ρ, Kruskal–Wallis H, тертильна стратифікація) та інтерпретації верифіковані авторами самостійно. ШІ-інструменти не використовувалися для генерації наукових тверджень, аналізу даних чи формулювання висновків.

Конфлікт інтересів

Автори заявляють про відсутність конфлікту інтересів та підтверджують, що під час підготовки цієї роботи не існувало жодних комерційних, фінансових чи інших взаємовідносин, які могли б бути розцінені як такі, що здатні вплинути на результати дослідження або їх інтерпретацію. Робота виконана відповідно до принципів академічної доброчесності, етичних норм проведення наукових досліджень та вимог редакційної політики щодо запобігання конфлікту інтересів.

Список використаної літератури:

1. Tawosi, V., Al-Subaihin, A., Moussa, R., & Sarro, F. (2022). A versatile dataset of agile open source software projects. In Proceedings of the 19th International Conference on Mining Software Repositories (MSR 2022) (pp. 707–711). https://doi.org/10.1145/3524842.3528029

2. Ясінецький, О. О., & Фесенко, Т. Г. (2025). Управління ризиками IT-проєктів: аналітичний огляд досліджень. Вісник Національного технічного університету «ХПІ». Серія: Стратегічне управління, управління портфелями, програмами та проєктами, (10), 98–109. https://doi.org/10.20998/2413-3000.2025.10.6

3. Ясінецький, О. О., & Фесенко, Т. Г. (2025). Інформаційна технологія оцінювання ризиків затримок у Scrum-проєктах IT-аутсорсингу на основі цифрових комунікаційних патернів і метрик продуктивності команд. Телекомунікаційні та інформаційні технології, (4), 160–167. https://doi.org/10.31673/2412-4338.2025.048919

4. Choetkiertikul, M., Dam, H. K., Tran, T., Pham, T., Ghose, A., & Menzies, T. (2019). A deep learning model for estimating story points. IEEE Transactions on Software Engineering, 45(7), 637–656. https://doi.org/10.1109/TSE.2018.2792473

5. Choetkiertikul, M., Dam, H. K., Tran, T., & Ghose, A. (2015). Predicting delays in software projects using networked classification. In 2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE) (pp. 353–364). https://doi.org/10.1109/ASE.2015.55

6. Kula, E., Greuter, E., Van Deursen, A., & Gousios, G. (2022). Factors affecting on-time delivery in large-scale agile software development. IEEE Transactions on Software Engineering, 48(9), 3573–3592. https://doi.org/10.1109/TSE.2021.3101192

7. Kula, E., Greuter, E., Van Deursen, A., & Gousios, G. (2023). Dynamic prediction of delays in software projects using delay patterns and Bayesian modeling. In Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2023) (pp. 1012–1023). https://doi.org/10.1145/3611643.3616328

8. Ortu, M., Adams, B., Destefanis, G., Tourani, P., Marchesi, M., & Tonelli, R. (2015). Are bullies more productive? Empirical study of affectiveness vs. issue fixing time. In Proceedings of the 12th Working Conference on Mining Software Repositories (MSR 2015) (pp. 304–313). https://doi.org/10.1109/MSR.2015.35

9. Schwaber, K., & Sutherland, J. (2020). The Scrum Guide: The definitive guide to Scrum: The rules of the game. https://scrumguides.org/scrum-guide.html

10. Project Management Institute. (2021). A guide to the project management body of knowledge (PMBOK Guide) (7th ed.). Project Management Institute.

11. Kitchenham, B., & Pfleeger, S. L. (2002). Principles of survey research. Part 5: Populations and samples. ACM SIGSOFT Software Engineering Notes, 27(5), 17–20. https://doi.org/10.1145/571681.571686

12. Stray, V., & Moe, N. B. (2020). Understanding coordination in global software engineering: A mixed-methods study on the use of meetings and Slack. Journal of Systems and Software, 170, 110717. https://doi.org/10.1016/j.jss.2020.110717

13. Kupiainen, E., Mäntylä, M. V., & Itkonen, J. (2015). Using metrics in Agile and Lean software development - A systematic literature review of industrial studies. Information and Software Technology, 62, 143–163. https://doi.org/10.1016/j.infsof.2015.02.005

References:

1. Tawosi, V., Al-Subaihin, A., Moussa, R., & Sarro, F. (2022). A versatile dataset of agile open source software projects. In Proceedings of the 19th International Conference on Mining Software Repositories (MSR 2022) (pp. 707–711). https://doi.org/10.1145/3524842.3528029

2. Yasinetskyi, O. O., & Fesenko, T. H. (2025). Upravlinnia ryzykamy IT-proiektiv: analitychnyi ohliad doslidzhen [Risk management of IT projects: An analytical review of research]. Visnyk Natsionalnoho Tekhnichnoho Universytetu «KhPI». Seriia: Stratehichne Upravlinnia, Upravlinnia Portfeliamy, Prohramamy ta Proiektamy, (10), 98–109. https://doi.org/10.20998/2413-3000.2025.10.6

3. Yasinetskyi, O. O., & Fesenko, T. H. (2025). Informatsiina tekhnolohiia otsiniuvannia ryzykiv zatrymok u Scrum-proiektakh IT-autsorsynhu na osnovi tsyfrovykh komunikatsiinykh paterniv i metryk produktyvnosti komand [Information technology for delay risk assessment in Scrum projects of IT outsourcing based on digital communication patterns and team productivity metrics]. Telekomunikatsiini ta Informatsiini Tekhnolohii, (4), 160–167. https://doi.org/10.31673/2412-4338.2025.048919

4. Choetkiertikul, M., Dam, H. K., Tran, T., Pham, T., Ghose, A., & Menzies, T. (2019). A deep learning model for estimating story points. IEEE Transactions on Software Engineering, 45(7), 637–656. https://doi.org/10.1109/TSE.2018.2792473

5. Choetkiertikul, M., Dam, H. K., Tran, T., & Ghose, A. (2015). Predicting delays in software projects using networked classification. In 2015 30th IEEE/ACM International Conference on Automated Software Engineering (ASE) (pp. 353–364). https://doi.org/10.1109/ASE.2015.55

6. Kula, E., Greuter, E., Van Deursen, A., & Gousios, G. (2022). Factors affecting on-time delivery in large-scale agile software development. IEEE Transactions on Software Engineering, 48(9), 3573–3592. https://doi.org/10.1109/TSE.2021.3101192

7. Kula, E., Greuter, E., Van Deursen, A., & Gousios, G. (2023). Dynamic prediction of delays in software projects using delay patterns and Bayesian modeling. In Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2023) (pp. 1012–1023). https://doi.org/10.1145/3611643.3616328

8. Ortu, M., Adams, B., Destefanis, G., Tourani, P., Marchesi, M., & Tonelli, R. (2015). Are bullies more productive? Empirical study of affectiveness vs. issue fixing time. In Proceedings of the 12th Working Conference on Mining Software Repositories (MSR 2015) (pp. 304–313). https://doi.org/10.1109/MSR.2015.35

9. Schwaber, K., & Sutherland, J. (2020). The Scrum Guide: The definitive guide to Scrum: The rules of the game. https://scrumguides.org/scrum-guide.html

10. Project Management Institute. (2021). A guide to the project management body of knowledge (PMBOK Guide) (7th ed.). Project Management Institute.

11. Kitchenham, B., & Pfleeger, S. L. (2002). Principles of survey research. Part 5: Populations and samples. ACM SIGSOFT Software Engineering Notes, 27(5), 17–20. https://doi.org/10.1145/571681.571686

12. Stray, V., & Moe, N. B. (2020). Understanding coordination in global software engineering: A mixed-methods study on the use of meetings and Slack. Journal of Systems and Software, 170, 110717. https://doi.org/10.1016/j.jss.2020.110717

13. Kupiainen, E., Mäntylä, M. V., & Itkonen, J. (2015). Using metrics in Agile and Lean software development - A systematic literature review of industrial studies. Information and Software Technology, 62, 143–163. https://doi.org/10.1016/j.infsof.2015.02.005

Надійшла до редакції: 28.02.26
Прийнята до друку:12.06.26
Опубліковано:30.06.26