

А. С. Сверстюк^{1,2}
Л. Є. Мосій¹

ПІДХІД ДО ОЦІНЮВАННЯ ЕЛЕКТРОКАРДІОСИГНАЛІВ НА ОСНОВІ БАГАТОФАКТОРНОГО РЕГРЕСІЙНОГО АНАЛІЗУ ФУНКЦІЇ ЧАСОВОЇ ВАРІАБЕЛЬНОСТІ

¹Тернопільський національний технічний університет імені І. Пулюя;

²Тернопільський національний медичний університет імені І. Я. Горбачевського

У дослідженні представлено розробку багатофакторної регресійної моделі, призначеної для автоматизованої ідентифікації кардіологічних станів шляхом аналізу статистичних параметрів часової мінливості електрокардіографічних даних. Розроблена модель базується на п'яти ключових статистичних індикаторах: середньому арифметичному, медіані, моді, середньоквадратичному відхиленні та ексцесі. Ці параметри ідентифіковано через послідовну процедуру регресійного відбору із загальної сукупності тринадцяти первинних характеристик за граничного рівня статистичної достовірності $p < 0,05$. Вагові параметри кожного індикатора визначено за допомогою оптимізації квадратичних відхилень. Домінантний вплив на результат діагностування чинить середнє арифметичне (коефіцієнт $\beta = 153,952$), натомість решта показників забезпечують уточнювальний ефект з від'ємними ваговими коефіцієнтами.

Розроблена модель демонструє високу ефективність розрізнення трьох категорій серцевої активності: фізіологічної норми, аритмічних розладів типу передчасних серцевих скорочень та структурних порушень провідності міокарда (часткової блокади лівого пучка Гіса). Якість моделювання підтверджується коефіцієнтом детермінації $R^2 = 0,97636$ (з урахуванням корекції $R^2 = 0,97577$). Фізіологічний стан характеризується мінімальною дисперсією часових інтервалів ($0,0003...0,00004$ с²), натомість передчасні скорочення демонструють тисячократне збільшення цього показника ($0,011...0,012$ с²). Перевірка на експериментальному масиві даних з 204 записів серцевої активності засвідчила високу надійність системи (показник Фішера $F = 1635,7$, коли $p < 0,001$) та дотримання фундаментальних вимог регресійного моделювання. Дослідження похибок прогнозування виявило їх відповідність гаусівському розподілу та однорідність дисперсії, що підтверджує коректність розробленої системи.

Запропонований підхід поєднує переваги класичних статистичних методів із запропонованим використанням функції часової варіабельності для комплексного аналізу морфологічних і ритмічних ознак кардіосигналів. Практичне значення дослідження полягає у створенні математичного інструментарію для автоматизованих систем діагностики серцево-судинних захворювань та систем підтримки прийняття клінічних рішень.

Ключові слова: функція часової варіабельності, багатофакторна регресійна модель, електрокардіосигнал, кардіологічна діагностика, покрокова регресія, екстрасистолія, блокада лівої ніжки пучка Гіса, коефіцієнт прогнозування CDPCTVF, статистичні предиктори, автоматична класифікація серцевих патологій, циклічні дискретні випадкові процеси, серцево-судинна система.

Вступ

Електрокардіографія (ЕКГ) є ключовою технологією в серцево-судинній медицині, яка забезпечує неінвазивне розуміння електричної активності, що лежить в основі функції серця. Форма хвилі ЕКГ фіксує послідовну деполяризацію та реполяризацію в провідній системі серця, відображаючи ритм серця та його електричну стабільність. Серед її основних компонентів інтервал RR — це час між послідовними зубцями R, та інтервал QT — тривалість від деполяризації шлуночків до реполяризації; служать ключовими показниками, які виражають часову варіабельність електричної активності серця. Ці варіабельності інтервалів відображають баланс між симпатичним та парасимпатичним тонуванням, здоров'ям міокарда та наявністю патофізіологічних процесів.

У статті [1] українськими вченими запропоновано новий алгоритм і метод отримання сигналів ЕКГ плода, який включає кілька етапів, зокрема, реєстрацію електрокардіосигналів (ЕКС) матері, синфазне виявлення ЕКС плода та адаптивну фільтрацію. Запропонований метод протестовано на згенерованому сигналі ЕКГ і встановлено, що він є ефективним для видалення ЕКС плода із сигналів, що містять шум та артефакти.

Підходи до моделювання, обробки та розпізнавання ЕКС можуть отримати суттєві переваги від впровадження теорії циклічних дискретних випадкових процесів (ЦДВП) [2], [3]. Ця методологія поєднує аналіз тимчасових та амплітудних характеристик, формуючи цілісну систему дослідження ЕКС з метою оцінювання функціонального стану серцево-судинної системи (ССС). Застосування окремих серцевих циклів як базових елементів аналізу створює умови для цільового дослідження з одночасним скороченням обсягу необхідних даних, спрощуючи ідентифікацію ключових часових маркерів та підвищуючи ефективність прогнозних алгоритмів [4].

Здатність досліджувати ЕКС в окремі часові точки забезпечує переваги у разі ідентифікації малопомітних порушень ритму, які залишаються непоміченими у разі застосування континуальних методів обробки, що має вирішальне значення для своєчасного виявлення та терапії кардіологічних патологій [5]. Поєднання амплітудного аналізу з часовими параметрами формує всебічну характеристику ЕКС, реєструючи одночасно інтенсивність та часові характеристики біоелектричної активності міокарда. Цифрова обробка сигналів гарантує отримання якісних результатів та забезпечує варіативність у разі архівування та подальшого дослідження [6].

Впровадження ЦДВП у практику дослідження ЕКС формує надійний методологічний фундамент для створення моделей, аналізу та діагностування характеристик серцевого ритму. Цей метод демонструє особливу цінність у разі дослідження як стабільних, так і аритмічних кардіоциклів. У порівнянні зі стандартними періодичними підходами, розроблена методологія забезпечує істотне підвищення точності, стабільності та діагностичної цінності комп'ютеризованого дослідження ЕКС. ЦДВП та їхні варіації, зокрема, циклічно корельовані стохастичні процеси, дозволяють математично формалізувати циклічну та випадкову природу ЕКС. Використання моделей на базі ЦДВП дає можливість адекватнішого відображення багатовимірних та фазових характеристик ЕКС, показуючи суттєві переваги над класичними періодичними підходами, зокрема у разі нестабільності серцевого ритму [7], [8].

Адаптовані до ритму статистичні оцінювальні процедури ґрунтуються на інваріантності параметрів ЦДВП відносно часових трансляцій. Ці алгоритми гарантують формування неупереджених та збіжних оцінок характеристик ЕКС. Практичні випробування виявляють істотне скорочення похибок порівняно з неадаптованими методиками, що набуває особливого значення для аналізу аритмічних записів [9]. Також використання ЦДВП дозволяє моделювати діагностичні простори, які допомагають у розмежуванні та групуванні діагностичних ознак. Це моделювання полегшує ідентифікацію груп спектральних коефіцієнтів, які відповідають нормальним та патологічним моделям ЕКС [10].

Алгоритми ЦДВП та ритм-орієнтовані методики ефективно імплементуються в спеціалізовані програмні комплекси для комп'ютеризованого дослідження ЕКС. Створені платформи виконують сегментування записів, визначення ритмічних характеристик та статистичну обробку в автоматизованому режимі. Технічні можливості цих комплексів охоплюють автоматичну ідентифікацію ритмічних порушень та інших відхилень, забезпечуючи підґрунтя для більш оперативної та точної діагностики кардіологічних захворювань [11], [12].

Аналіз підходів до оцінювання ЕКС на основі багатфакторного регресійного аналізу

Вимірювання ЕКС за допомогою багатовимірної регресійної моделі для характеристики часової варіабельності ритму передбачає інтеграцію передових статистичних моделей з класичними та новими електрофізіологічними ознаками ЕКС. З огляду на складну та взаємозалежну природу ЕКС та їхню модуляцію клінічними факторами, однофакторний аналіз часто не враховує повністю основні коваріаційні структури та впливи, що впливають на результат. Методи багатфакторної регресії пропонують надійну статистичну основу для одночасного моделювання кількох, потенційно корельованих ознак ЕКС разом з екзогенними коваріатами, такими як вік, стать, статус захворювання або фармакологічні втручання. Такі моделі враховують складні взаємозв'язки, зокрема, лінійні та нелінійні залежності, що дозволяє розкласти дисперсію, яка пояснюється різними фізіологічними та патологічними факторами, одночасно підвищуючи точність прогнозування.

Цей підхід відрізняється від простої лінійної регресії або одновимірного аналізу, які можуть не враховувати важливі взаємодії або призводити до оманливих висновків, коли змінні моделюються незалежно. Непараметричні методи регресії, такі як багатовимірні адаптивні регресійні сплайни (MARS), ще більше розширюють цю парадигму, фіксуючи нелінійні закономірності без заздалегідь визначених функціональних форм, підвищуючи гнучкість моделі та продуктивність ЕКС-візуалізації та оцінки часу активації. До того ж персоналізовані регресійні підходи, що враховують ідіосинкратичні співвідношення QT/RR, ілюструють необхідність моделювання індивідуально-специфічної динаміки, а не покладатися на середні показники популяції [13]—[15].

Методології багатовимірної регресії в аналізі ЕКС використовуються для досягнення численних цілей у клінічних та дослідницьких сферах. Насамперед ці моделі розроблені для прогнозування серцево-судинних патологій та клінічних результатів, зокрема, прогресування серцевої недостатності, наявності аритмій та несприятливих серцевих подій. Інтегруючи часові, спектральні та морфологічні ознаки ЕКС зі специфічними для пацієнта коваріатами, ці моделі можуть стратифікувати пацієнтів точніше ніж однофакторні предиктори.

Характеристика індивідуальної мінливості є ще однією критично важливою метою. Враховуючи, що співвідношення інтервалів ЕКС, такі як зв'язок QT/RR, демонструють суб'єктивно-специфічні закономірності, багатовимірна регресія надає механізм моделювання цієї гетерогенності, тим самим підвищуючи діагностичну точність та персоналізоване застосування в медицині. Більше того, часова динаміка є фундаментальною, оскільки серцевий ритм змінюється протягом циркадних циклів, станів сну та у відповідь на екологічні чи патологічні подразники. Застосування охоплює все від оцінки серцево-судинних патологій (ССП) та класифікації серцебиття в режимі реального часу. Цей широкий спектр підтверджується досягненнями в обчислювальній потужності, обробці сигналів та машинному навчанні, які разом сприяють багатовимірному аналізу, який поєднує електрофізіологічну теорію з прикладною клінічною допомогою [13], [16], [17].

Окрім часових інтервалів, частотні характеристики надають уявлення про вегетативну модуляцію серцевого ритму. Спектральний аналіз зазвичай поділяє мінливість на низькочастотні (НЧ) та високочастотні (ВЧ) діапазони, які зазвичай інтерпретуються як показники симпатичної та парасимпатичної (вагальної) активності відповідно. Співвідношення НЧ/ВЧ широко використовується як маркер симпатовагального балансу. Обчислювальні підходи включають авторегресивні моделі для оцінки спектральної щільності потужності, тригонометричні та вейвлет-розклади для розв'язання періодичних складових, а також аналіз передавальних функцій для оцінки чутливості барорефлексу.

Проте ці методи значною мірою залежать від припущень про стаціонарність сигналу, які можуть бути порушені в амбулаторних або динамічних умовах, і чутливі до шуму та артефактів. Отже, методологічні вдосконалення та протоколи контрольованого збору даних є важливими для забезпечення валідності. Інтеграція спектральних характеристик з часом та нелінійними показниками підвищує інтерпретаційну здатність багатовимірних регресійних моделей [18]—[20].

Моделі машинного навчання, включно з рекурентними нейронними мережами з довгостроковою пам'яттю (LSTM), використовувалися для регресійного прогнозування характеристик серцевого ритму з даних ЕКС часових рядів. Архітектура LSTM, що вміє фіксувати часові залежності та керувати вхідними даними змінної довжини, особливо добре підходить для прогнозування варіабельності серцевого ритму та класифікації типів аритмій в умовах високої варіабельності ритму, такої як синусова тахікардія. Протоколи вибору ознак, зокрема, мінімальна надмірність та максимальна релевантність, визначають найінформативніші ознаки HRV у часовій та частотній областях для оптимізації навчання моделі. Ці моделі досягають високої точності прогнозування та валідації, демонструючи доцільність застосування в реальному часі та амбулаторних умовах [21], [22].

Інтеграція багатовимірної регресії з машинним навчанням та обчислювальним моделюванням призвела до суттєвих покращень у неінвазивному моніторингу ЕКС, класифікації ритму та безперервному оцінюванні здоров'я в різних популяціях [24], [25]. Ці методи застосовано до носимих технологій, моніторингу фізичних вправ у режимі реального часу, виявлення хронічного стресу та досліджень раннього розвитку новонароджених, значно розширивши їхнє клінічне охоплення [17], [22], [26].

Незважаючи на досягнення, багатовимірні регресійні моделі варіабельності ЕКС стикаються з обмеженнями, включаючи чутливість до шуму сигналу, нестационарність фізіологічних записів та наявність артефактів. Індивідуальна фізіологічна варіабельність ще більше ускладнює зусилля щодо формулювання універсальних прогностичних моделей. Забруднення артефактами, особливо

в амбулаторних ЕКГ, може спотворювати показники ВСР та результати регресійної помилки. До того ж розбіжності в протоколах збору даних та обмежені розміри вибірки в багатьох дослідженнях перешкоджають узагальненню. Ці проблеми вимагають постійних методологічних інновацій [15], [23], [24]. Розширення застосування багатовимірної регресії в розвитку кардіології, оцінюванні стресу та когнітивного навантаження, а також на платформи носимих датчиків ще більше збагачує розуміння та лікування серцево-судинних захворювань [13], [17], [25].

Метою роботи є створення та експериментальна перевірка багатофакторної регресійної моделі, здатної автоматизовано ідентифікувати кардіологічні стани шляхом аналізу статистичних параметрів часової варіабельності електрокардіографічних даних. Розроблювана модель має забезпечувати неупереджену числову характеристику роботи ССС та надійне розмежування фізіологічних і патологічних варіантів серцевої активності.

Модель функції часової варіабельності з урахуванням екстремальних амплітудних значень характеристичних зубців ЕКС

У роботі [3] запропонована математична модель функції часової варіабельності (ФЧВ) з урахуванням екстремальних амплітудних значень характеристичних зубців ЕКС для комплексного аналізу морфологічних і ритмічних діагностичних ознак кардіосигналів. Ця модель представлена функцією $T_{A_k}(m)$, яка враховує екстремальні амплітудні значення характеристичних зубців ЕКС (P , Q , R , S і T).

$$T_{A_k}(m) = t_{A_k}(m) - t_{A_k}(m-1), \quad k \in \{P, Q, R, S, T\}, \quad m \in \mathbf{Z}, \quad (1)$$

де $t_{A_k}(m)$ — момент часу досягнення піку k -го типу зубця в m -му кардіоциклі, с; $t_{A_k}(m-1)$ — момент часу досягнення піку k -го типу зубця в попередньому кардіоциклі ($m-1$), с; $T_{A_k}(m)$ — значення ФЧВ з урахуванням екстремальних амплітудних значень характеристичних зубців ЕКС, що відображає часовий інтервал між піками зубців k -го типу у поточному m та попередньому кардіоциклі ($m-1$); $k \in \{P, Q, R, S, T\}$ — тип характеристичного зубця; $m \in \mathbf{Z}$ — номери циклів.

Функція ФЧВ з урахуванням екстремальних амплітудних значень характеристичних зубців ЕКС $T_{A_k}(m)$ характеризується такими властивостями:

- визначена для всіх $m \geq 2$, оскільки потребує попереднього циклу для обчислення.
- область значень: $T_{A_k}(m) \in (0, +\infty)$, с.

Для кількісного опису функції $T_{A_k}(m)$ використано метод статистичної обробки, що дозволяє розрахувати такі статистичні параметри:

- оцінка математичного сподівання часових інтервалів

$$\hat{m}_{T_{A_k}} = \frac{1}{M} \sum_{m=1}^M T_{A_k}(m) = \sum_{m=1}^M [t_{A_k}(m) - t_{A_k}(m-1)]; \quad (2)$$

- оцінка дисперсії часових інтервалів

$$\hat{d}_{T_{A_k}} = \frac{1}{M-1} \sum_{m=1}^M [T_{A_k}(m) - \hat{m}_{T_{A_k}}]^2 = \frac{1}{M-1} \sum_{m=1}^M [(t_{A_k}(m) - t_{A_k}(m-1)) - \hat{m}_{T_{A_k}}]^2; \quad (3)$$

- розмах значень (варіаційний розмах) часових інтервалів

$$R_{T_{A_k}} = \max_{m=1, 2, \dots, M} T_{A_k}(m) - \min_{m=1, 2, \dots, M} T_{A_k}(m), \quad (4)$$

де M — загальна кількість аналізованих кардіоциклів.

Проведені дослідження продемонстрували перспективність застосування цієї функції для комплексного аналізу морфологічних і ритмічних діагностичних ознак ЕКС. Зокрема, встановлено суттєві відмінності статистичних характеристик ФЧВ між станами умовної норми та патологічними змінами серцевого ритму.

У пацієнта з фізіологічною нормою серцевої діяльності зафіксовано високий рівень стабільності часових проміжків між основними елементами ЕКС. Математичне сподівання тривалості інтервалів для характерних хвиль (P , R , T) дорівнювало 0,776 с, що є індикатором ритмічності серцевих

скорочень. Показник дисперсії коливався в діапазоні $0,00003...0,00004$ с², при цьому різниця між максимальним і мінімальним значеннями становила від $0,016$ до $0,024$ с. Ці параметри демонструють незначну мінливість тривалості між сусідніми серцевими циклами. А для пацієнта з діагностованою екстрасистолією характерні істотні порушення ритмічності серцевих скорочень. Середня тривалість часової функції серцевого ритму скоротилася до діапазону $0,503...0,504$ с, демонструючи зниження на 35 % у порівнянні з показниками фізіологічної норми. Величина дисперсії збільшилася в тисячу разів, досягнувши значень $0,011...0,012$ с². Амплітуда коливань часових інтервалів становила $0,408...0,496$ с, перевищуючи нормальні значення в 17—31 рази. Виявлені відхилення є характерними ознаками аритмічності серцевої діяльності у разі екстра-систолічних порушень [3].

Встановлені суттєві розбіжності у статистичних характеристиках ФЧВ під час порівняння фізіологічної норми та патологічних станів вказують на значний діагностичний потенціал розробленого підходу. Проте для впровадження ФЧВ у медичну практику потрібно створити математичний апарат для автоматизованого визначення стану хворого, що базується на розрахованих характеристиках. Зважаючи на комплексний характер часової мінливості, раціональним рішенням є використання багатопараметричного регресійного моделювання.

Результати дослідження

З метою створення та перевірки багатфакторної регресійної моделі прогнозування кардіологічного діагнозу (БРМПКД) з використанням ФЧВ підготовлено статистично значущу базу даних з 204 ЕКС. Розподіл досліджуваного матеріалу такий:

- 102 записи (50 %) — ЕКС осіб без виражених порушень функціонування ССС, з яких 51 запис від пацієнтів з імплантованим електрокардіостимулятором та 51 — без такого пристрою;
- 51 запис (25 %) — дані хворих з аритміями серця, що проявлялися у вигляді різноманітних типів екстрасистолічних порушень;
- 51 запис (25 %) — ЕКС осіб зі структурними змінами провідної системи серця, зокрема неповною блокадою лівої ніжки пучка Гіса.

Описаний розподіл матеріалу гарантує рівномірне охоплення ключових видів серцевої патології та створює умови для формування універсального класифікаційного алгоритму.

На початковому етапі аналізу проводилося визначення набору з 13 статистичних параметрів для кожного масиву даних ФЧВ:

1. Середнє арифметичне — відображає основну тенденцію розподілу часових варіацій.
2. Стандартна похибка середнього — визначає надійність оцінки центрального значення.
3. Медіана — стійкий показник центру розподілу, нечутливий до аномальних даних.
4. Мода — найчастіше спостережуване значення часових варіацій.
5. Середньоквадратичне відхилення (СКВ) — описує ступінь розкиду даних відносно центру.
6. Дисперсія вибірки — квадратична міра мінливості значень.
7. Коефіцієнт ексцесу — визначає ступінь концентрації даних навколо центру розподілу.
8. Коефіцієнт асиметрії — характеризує нерівномірність розподілу.
9. Діапазон варіювання — відстань між крайніми значеннями вибірки.
10. Мінімальне значення — найменший показник часових варіацій.
11. Максимальне значення — найбільший показник часових варіацій.
12. Сумарний показник — загальна сума компонентів вибірки.
13. 95 %-й довірчий рівень — границя інтервалу, що містить справжнє середнє з вірогідністю 95 %.

Для створення БРМПКД використовувався алгоритм поетапної регресії з послідовним включенням незалежних змінних. Аналіз експериментальних даних здійснювався за допомогою спеціалізованого програмного забезпечення Statistica 12.0 (виробництва StatSoft Inc., США). Для додавання факторів до регресійної моделі визначено поріг статистичної значущості $p < 0,05$, тоді як для їх видалення застосовувався критерій $p > 0,10$. Використання методики поетапного регресійного моделювання надало можливість відібрати з первинного переліку 13-ти статистичних характеристик 5 найінформативніших параметрів ($p < 0,001$), які сформували остаточну модель: середнє арифметичне, медіана, мода, СКВ та ексцес.

Дослідження коефіцієнтів регресії (рис. 1) уможливило визначення вкладу окремих незалежних змінних у розрахунок прогнозного показника.

Regression Summary for Dependent Variable: CDPCTVF (1 in CDPCTVF)						
R= ,98811081 R²= ,97636298 Adjusted R²= ,97576609 F(5, 198)=1635,7 p<0,0000 Std. Error of estimate: ,17448						
N=204	b*	Std. Err. of b*	b	Std. Err. of b	t(198)	p-value
Intercept			-22,723	0,40650	-55,8992	0,000000
Mean	14,3021	1,083756	153,952	11,66584	13,1968	0,000000
Median	-10,6606	1,085992	-113,064	11,51775	-9,8165	0,000000
Mode	-0,3499	0,030347	-3,178	0,27566	-11,5292	0,000000
StDev	-3,7799	0,048691	-52,790	0,68002	-77,6305	0,000000
Kurtosis	-0,9002	0,025099	-0,988	0,02756	-35,8643	0,000000

Рис. 1. Результат значущих факторів для коефіцієнта прогнозування кардіологічного діагнозу ФЧВ для проведення багатофакторного регресійного аналізу в програмі Statistica 12.0

Найвищий додатний коефіцієнт спостерігається для параметра середнє арифметичне ($\beta = 153,952$), що свідчить про його визначальну роль у формуванні діагностичного висновку. Від'ємні значення коефіцієнтів для медіани ($\beta = -113,064$), СКВ ($\beta = -52,790$), моди ($\beta = -3,178$) та ексцесу ($\beta = -0,988$) демонструють обернену залежність між цими характеристиками та діагностичним показником. Вільний член моделі становить $-22,723$. Нульові значення p -рівня ($p = 0,000000$) пов'язані з технічними обмеженнями відображення в програмі Statistica 12.0. Реальні показники значущості для всіх факторів менше 0,001, при цьому для ключових незалежних змінних вони досягають величин 10^{-15} .

Використовуючи результати регресійного моделювання, розроблено математичну модель для визначення коефіцієнта прогнозування кардіологічного діагнозу функції часової варіабельності (CDPCTVF – Cardiac Diagnosis Prediction Coefficient of the Time Variability Function):

$$\text{CDPCTVF} = 153,952 \cdot \text{Середнє} - 113,064 \cdot \text{Медіана} - 3,178 \cdot \text{Мода} - 52,790 \cdot \text{СКВ} - 0,988 \cdot \text{Ексцес} - 22,723.$$

Усі змінні обчислюються для індивідуального масиву ФЧВ обстежуваного хворого.

Фундаментальним кроком верифікації регресійної моделі виступає ретельне дослідження залишків. Частотний розподіл залишкових величин (рис. 2) показує симетричну конфігурацію з вираженою центральною модою та рівномірним спаданням частот до країв розподілу. Аналіз форми розподілу вказує на його відповідність теоретичному нормальному закону, що забезпечує коректність використання методології найменших квадратів.

З метою підтвердження нормальності залишків побудовано нормально-ймовірнісний графік (Q-Q plot) (рис. 3). Дослідження графіка показало розміщення емпіричних даних уздовж теоретичної лінії без систематичних зміщень, що свідчить про нормальний характер розподілу залишків. Мінімальні розбіжності в зонах крайніх значень знаходяться в межах статистичної похибки та не порушують загальну валідність моделі.

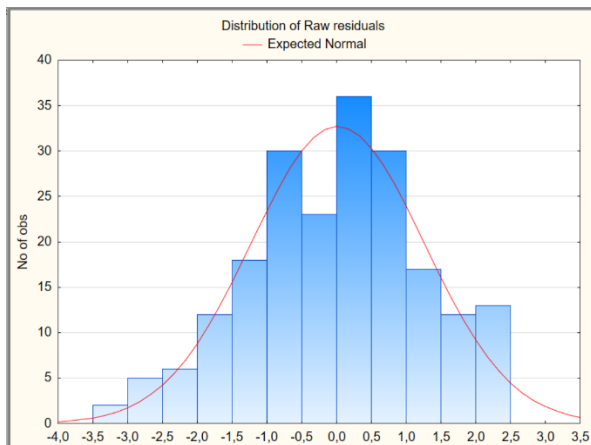


Рис. 2. Гістограма залишкових відхилень багатофакторної регресійної моделі прогнозування кардіологічного діагнозу на основі функції часової варіабельності

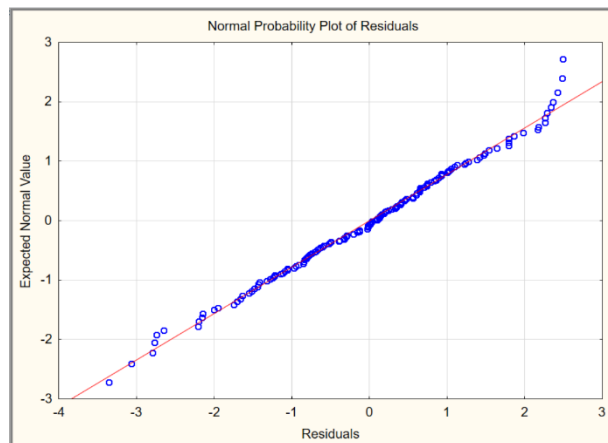


Рис. 3. Нормально-ймовірнісний графік залишкових відхилень багатофакторної регресійної моделі прогнозування кардіологічного діагнозу на основі функції часової варіабельності

Ключовою умовою регресійного моделювання виступає сталість варіації залишків (гомоскедастичність). Перевірка цієї умови здійснювалася через побудову графіка розкиду залишків щодо передбачених величин (рис. 4). Аналіз графічного зображення виявляє випадковий характер розміщення даних без ознак систематичної зміни варіації залежно від рівня прогнозу. Відсутність

типових артефактів (конусоподібного розширення, нелінійних залежностей) підтверджує дотримання принципу гомоскедастичності.

Перевірка загальної статистичної достовірності моделі здійснювалася через дисперсійний аналіз (ANOVA) (рис. 5). Отримані дані підтверджують високу достовірність регресійної залежності: F-статистика дорівнює 1635,7 з рівнем значущості $p < 0,001$. Це доводить, що побудована модель забезпечує достовірне прогнозування в порівнянні з нульовою гіпотезою (використання тільки середнього).

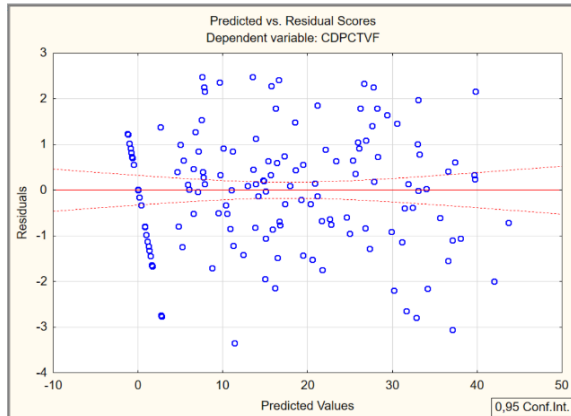


Рис. 4. Діаграма розсіювання залишкових відхилень регресійної моделі прогнозування кардіологічного діагнозу на основі функції часової варіабельності

Analysis of Variance; DV: CDPCTVF (1 in CDPCTVF)					
Effect	Sums of Squares	df	Mean Squares	F	p-value
Regress.	25982,23	15	1732,149	1005,241	0,00
Residual	323,95	188	1,723		
Total	26306,18				

Рис. 5. Результати аналізу ANOVA

Показник детермінації Нейджелкерка є комплексним критерієм ефективності регресійної залежності, що відображає частку поясненої варіації цільової змінної за рахунок вибраних факторів. Для створеної моделі $R^2 = 0,97636298$ (скоригований $R^2 = 0,97576609$), що відповідає поясненню 97,6 % мінливості діагностичного показника вибраними статистичними параметрами ФЧВ. Настільки значний коефіцієнт детермінації вказує на правильний відбір незалежних змінних для побудови моделі, відсутність істотних неврахованих чинників, потужні прогностичні можливості моделі, придатність для впровадження в автоматизовані діагностичні системи.

Розроблена математична модель забезпечує неупереджене кількісне оцінювання функціонального статусу ССС через дослідження часової мінливості ЕКС. Підвищені показники коефіцієнта CDPCTVF вказують на патологічні процеси, натомість знижені величини відповідають нормальній серцевій діяльності.

Вивчення компонентів моделі показало, що визначальний вплив на діагностичний результат має показник середнього арифметичного, який віддзеркалює базовий рівень часової мінливості. Решта характеристик (медіана, мода, СКВ, ексцес) виконують коригувальну функцію, деталізуючи діагностичний висновок через врахування різних аспектів розподілу часових інтервалів.

Висновки

За підсумками проведеного дослідження розроблена та перевірена багатофакторна регресійна модель прогнозування кардіологічного діагнозу (CDPCTVF) на основі ФЧВ ЕКС. Модель забезпечує високоточну кількісну оцінку функціонального стану ССС, що відображає коефіцієнт детермінації $R^2 = 0,97636$ (скоригований $R^2 = 0,97577$), демонструючи врахування 97,6 % варіації діагностичного коефіцієнта вибраними статистичними показниками.

Застосування методу покрокової регресії дозволило з початкового набору 13 статистичних характеристик ФЧВ виділити 5 статистично значущих ($p < 0,001$): середнє арифметичне, медіану, моду, СКВ та ексцес. Найбільший внесок у формування діагностичного висновку має показник середнього арифметичного ($\beta = 153,952$), тоді як медіана ($\beta = -113,064$), СКВ ($\beta = -52,790$), мода ($\beta = -3,178$) та ексцес ($\beta = -0,988$) виконують коригувальну функцію через від'ємні коефіцієнти регресії.

Комплексна статистична перевірка підтвердила високу надійність та адекватність розробленої моделі. F-критерій Фішера становить 1635,7 з рівнем значущості $p < 0,001$, що свідчить про статистичну достовірність регресійної залежності. Аналіз залишкових відхилень продемонстрував їх відповідність нормальному закону розподілу та виконання умови гомоскедастичності, що підтвер-

джує коректність вибраної лінійної структури моделі та відсутність систематичних похибок прогнозування. Запропонований підхід успішно поєднує переваги класичних регресійних моделей з інноваційним використанням ФЧВ для комплексного аналізу морфологічних і ритмічних діагностичних ознак ЕКС.

Перспективи подальших досліджень охоплюють розширення спектру диференційованих патологій ССС, інтеграцію з гібридними підходами машинного навчання для підвищення робастності у разі роботи із зашумленими сигналами, а також проведення розширеної клінічної апробації на репрезентативних вибірках пацієнтів різних вікових груп та з різною кардіологічною патологією.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] H. Franchevska, M. Khvostivskiy, V. Dozorskiy, E. Yavorska, and O. Zastavnyy, "The Method and Algorithm for Detecting the Fetal ECG Signal in the Presence of Interference," in *Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023)*, CEUR Workshop Proceedings, Ternopil, Ukraine, June 14-16, 2023, pp. 263-272. ISSN 1613-0073.
- [2] L. Mosiy, and A. Sverstiuk, "Methods of modeling and classification of electrocardiograms," *Computer-Integrated Technologies: Education, Science, Production*, no. 58, pp. 104-115, 2025. <https://doi.org/10.36910/6775-2524-0560-2025-58-12>.
- [3] L. Mosiy, and A. Sverstiuk, "Information technology for electrocardiographic signal analysis based on mathematical models of temporal and amplitude variability," *Computer Systems and Information Technologies*, no. 2, pp. 36-44, 2025. <https://doi.org/10.31891/csit-2025-2-4>.
- [4] Y. Kim, et al., "Predicting Future Incidences of Cardiac Arrhythmias Using Discrete Heartbeats from Normal Sinus Rhythm ECG Signals via Deep Learning Methods," *Diagnostics*, vol. 13, no. 17, p. 2849, 2023. <https://doi.org/10.3390/diagnostics13172849>.
- [5] P. K. P. Shanmuganathan, and V. Sivaratri, "Integrating Deep Learning for Arrhythmia Detection with Automated Drug Delivery: A Comprehensive Approach to Cardiac Health Monitoring and Treatment," *arXiv*, 2024. <https://doi.org/10.48550/ARXIV.2410.19827>.
- [6] R. Patil, B. Narkhede, S. Varma, S. Suraliya, and N. Mehendale, "Auto Lead Extraction and Digitization of ECG Paper Records using cGAN," *arXiv*, 2022. <https://doi.org/10.48550/ARXIV.2211.06720>.
- [7] S. Lupenko, and R. Butsiy, "Isomorphic Multidimensional Structures of the Cyclic Random Process in Problems of Modeling Cyclic Signals with Regular and Irregular Rhythms," *Fractal and Fractional*, vol. 8, no. 4, p. 203, 2024. <https://doi.org/10.3390/fractalfract8040203>.
- [8] S. Lupenko, "The Mathematical Model of Cyclic Signals in Dynamic Systems as a Cyclically Correlated Random Process," *Mathematics*, vol. 10, no. 18, pp. 3406, 2022. <https://doi.org/10.3390/math10183406>.
- [9] S. Lupenko, "Rhythm-adaptive statistical estimation methods of probabilistic characteristics of cyclic random processes," *Digital Signal Processing*, vol. 151, pp. 104563, 2024. <https://doi.org/10.1016/j.dsp.2024.104563>.
- [10] V. Martsenyuk, A. Sverstiuk, A. Klos-Witkowska, A. Horkunenko, and S. Rajba, "Vector of Diagnostic Features in the Form of Decomposition Coefficients of Statistical Estimates Using a Cyclic Random Process Model of Cardiosignal," in *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, Metz, France, 2019, pp. 298-303. <https://doi.org/10.1109/idaacs.2019.8924398>.
- [11] S. Lupenko, I. Lytvynenko, A. Sverstiuk, B. Shelestovskiy, and A. Horkunenko, "Software for Statistical Processing and Modeling of a Set of Synchronously Registered Cardio Signals of Different Physical Nature," *CEUR Workshop Proceedings*, vol. 2864, pp. 194-205, 2021. <https://doi.org/10.32782/cmws/2864-17>.
- [12] V. Kalidas and L. S. Tamil, "Detection of atrial fibrillation using discrete-state Markov models and Random Forests," *Computers in Biology and Medicine*, vol. 113, pp. 103386, 2019. <https://doi.org/10.1016/j.combiomed.2019.103386>.
- [13] T. K. Lin, D. Wendelken, B. O'Rourke, S. R. Jones, N. M. Punjabi, and D. DeMazumder, "Abstract 453: Phenotyping by ECG Dynamics During Sleep Predicts Cardiovascular Disease Risk Factors in a Multicenter Study of Asymptomatic Middle-aged Community Adults," *Circulation Research*, vol. 125, Suppl. 1, 2019. https://doi.org/10.1161/res.125.suppl_1.453.
- [14] A. Mombereau, Y. Serinagaoglu Dogrusoz, R. Dubois, and L. R. Bear, "Improved Performance of Data-Adaptive Regression Framework Based on Multivariate Adaptive Regression Splines for Electrocardiographic Imaging," in *Computing in Cardiology Conference (CinC)*, vol. 51, 2024. <https://doi.org/10.22489/cinc.2024.294>.
- [15] M. Malik, "Relation between QT and RR intervals is highly individual among healthy subjects: implications for heart rate correction of the QT interval," *Heart*, vol. 87, no. 3, pp. 2200-228, 2002. <https://doi.org/10.1136/heart.87.3.220>.
- [16] A. Paul, et al., "Development of Automated Cardiac Arrhythmia Detection Methods Using Single Channel ECG Signal," 2023.
- [17] M. Lavanga, et al., "Maturation of the Autonomic Nervous System in Premature Infants: Estimating Development Based on Heart-Rate Variability Analysis," *Frontiers in Physiology*, vol. 11, 2021. <https://doi.org/10.3389/fphys.2020.581250>.
- [18] B. A. Shah, et al., "Heart rate variability and amplitude-integrated electroencephalography measured shortly after birth and time to reach clinical milestones: a pilot study in late preterm infants," *Frontiers in Pediatrics*, vol. 13, 2025. <https://doi.org/10.3389/fped.2025.1579197>.
- [19] R. Sassi, et al., "Advances in heart rate variability signal analysis: joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society," *Europace*, vol. 17, no. 9, pp. 1341-1353, 2015. <https://doi.org/10.1093/europace/euv015>.
- [20] M. Pagani, D. Lucini, and A. Porta, "Sympathovagal balance from heart rate variability: time for a second round?" *Experimental Physiology*, vol. 97, no. 10, pp. 1141-1142, 2012. <https://doi.org/10.1113/expphysiol.2012.066977>.
- [21] B. Dhananjay, and J. Sivaraman, "Prediction of cardiac rhythm based on heart rate variability features in sinus tachycardia conditions," in *2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, 2021, pp. 1-6. <https://doi.org/10.1109/icbsii51839.2021.9445167>.

[22] A. Smiley and J. Finkelstein, "Regression Analysis for Prediction of Exercise Exertion Levels Using Physiological Data," in *2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, USA, 2023, pp. 0416-0419. <https://doi.org/10.1109/uemcon59035.2023.10315969>.

[23] C. D. Zuluaga-Ríos, M. A. Álvarez-López, and Á. A. Orozco-Gutiérrez, "Comparación de métodos de filtro de Kalman robustos para corrección de artefactos en el análisis de la variabilidad de la frecuencia cardíaca," *Tecnológicas*, vol. 18, no. 34, p. 25, 2015. <https://doi.org/10.22430/22565337.213>.

[24] P. Sardar, R. Gupta, and S. Mukhopadhyay, "Multiclass Signal Quality Assessment of Electrocardiogram using Entropy-based Features and Machine Learning Technique," in *2022 IEEE Silchar Subsection Conference (SILCON)*, Silchar, India, 2022, pp. 1-7. <https://doi.org/10.1109/silcon55242.2022.10028787>.

[25] A. Lyon, A. Mincholé, J. P. Martínez, P. Laguna, and B. Rodriguez, "Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances," *Journal of The Royal Society Interface*, vol. 15, no. 138, pp. 20170821, 2018. <https://doi.org/10.1098/rsif.2017.0821>.

[26] J. Li, M. Wang, F. Zhang, G. Liu, and W. Wen, "Chronic Stress Recognition Based on Time-Slot Analysis of Ambulatory Electrocardiogram and Tri-Axial Acceleration," *IEEE Transactions on Affective Computing*, vol. 15, no. 3, pp. 1178-1189, 2024. <https://doi.org/10.1109/taffc.2023.3326747>.

Рекомендована кафедрою біомедичної інженерії та оптико-електронних систем ВНТУ

Стаття надійшла до редакції 12.09.2025

Сверстюк Андрій Степанович — д-р техн. наук, професор, професор кафедри медичної інформатики Тернопільського національного медичного університету імені І. Я. Горбачевського; професор кафедри комп'ютерних наук Тернопільського національного технічного університету імені Івана Пулюя, e-mail: sverstyuk@tdmu.edu.ua;

Мосій Любомир Євгенійович — аспірант кафедри комп'ютерних наук, e-mail: lmosiy@gmail.com.

Тернопільський національний технічний університет імені Івана Пулюя, Тернопіль

A. S. Sverstiuk^{1,2}

L. Ye. Mosiy¹

Approach to Electrocardiographic Signal Assessment Based on Multifactor Regression Analysis of Time Variability Function

¹Ternopil Ivan Puluj National Technical University;

²Ternopil Ivan Horbachevsky National Medical University

The study presents the development of a multifactor regression model designed for automated identification of cardiac conditions through the analysis of statistical parameters of electrocardiographic data temporal variability. The developed model is based on five key statistical indicators: arithmetic mean, median, mode, root mean square deviation, and kurtosis. These parameters were identified through a sequential regression selection procedure from a total set of thirteen primary characteristics at a threshold level of statistical significance $p < 0.05$. The weight parameters of each indicator were calculated using quadratic deviation optimization. The arithmetic mean demonstrates the dominant influence on the diagnostic outcome (coefficient $\beta = 153.952$), while the remaining indicators provide a refining effect with negative weight coefficients.

The developed model demonstrates high efficiency in distinguishing three categories of cardiac activity: physiological norm, arrhythmic disorders such as premature cardiac contractions, and structural myocardial conduction disturbances (partial left bundle branch block). The modeling quality is confirmed by the coefficient of determination $R^2 = 0.97636$ (adjusted $R^2 = 0.97577$). The physiological state is characterized by minimal variance of time intervals ($0.00003\text{--}0.00004\text{ s}^2$), while premature contractions demonstrate a thousand-fold increase in this indicator ($0.011\text{--}0.012\text{ s}^2$). Validation on an experimental dataset of 204 cardiac activity records confirmed the high reliability of the system (Fisher's F -statistic = 1635.7 at $p < 0.001$) and compliance with fundamental requirements of regression modeling. The analysis of prediction errors revealed their correspondence to Gaussian distribution and variance homogeneity, confirming the correctness of the developed system.

The proposed approach combines the advantages of classical statistical methods with the proposed use of the time variability function for comprehensive analysis of morphological and rhythmic characteristics of cardiac signals. The practical significance lies in creating mathematical tools for automated cardiovascular disease diagnostic systems and clinical decision support systems.

Keywords: time variability function, multifactor regression model, electrocardiographic signal, cardiac diagnosis, stepwise regression, extrasystole, left bundle branch block, CDPCTVF prediction coefficient, statistical predictors, automatic classification of cardiac pathologies, cyclic discrete random processes, cardiovascular system.

Sverstiuk Andriy S. — Dr. Sc. (Eng.), Professor, Professor of the Chair of Medical Informatics of Ternopil Ivan Horbachevsky National Medical University; Professor of the Chair of Computer Science of Ternopil Ivan Puluj National Technical University, e-mail: sverstyuk@tdmu.edu.ua;

Mosiy Lyubomyr Ye. — Post-Graduate Student of the Chair of Computer Science, e-mail: lmosiy@gmail.com