

# Вибір оптимального порядку мультиваріативних авторегресійних моделей електроенцефалограм при епілепсії

Котючий І. В.<sup>1</sup>, Попов А. О.<sup>1</sup>, Харитонов В. І.<sup>2</sup>

<sup>1</sup>Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського“

<sup>2</sup>ТМО “Психіатрія”, м. Київ

E-mail: ivanellokot@gmail.com

В роботі запропонований підхід до вибору оптимальних порядків мультиваріативних авторегресійних (МВАР) моделей сигналів електричної активності мозку пацієнтів, яким поставлений діагноз епілепсії. Проведено МВАР-моделювання та статистичний аналіз оптимальних порядків моделей ділянок сигналів до, під час та після епілептичного нападу. Для оцінки якості моделей використано інформаційний критерій Шварца-Байеса з урахуванням коваріаційної матриці залишків змодельованих даних. Додатково якість оцінено за допомогою коефіцієнта кореляції Пірсона між реальними та змодельованими даними. В експериментальній частині були використані записи електроенцефалографічних (ЕЕГ) сигналів з двох наборів даних для генералізованих та фокальних епілептичних нападів. За результатами проведеної роботи рекомендується обирати значення порядків МВАР моделей ЕЕГ-сигналів для зазначених ділянок в діапазоні 11-13. Вказані шляхи подальших досліджень електричної активності мозку та функціональних зв'язків його відділів під час епілептичної активності.

*Ключові слова:* авторегресійна модель; електроенцефалографія; епілепсія; порядок статистичної моделі; епілептичні напади

DOI: [10.20535/RADAP.2018.73.33-39](https://doi.org/10.20535/RADAP.2018.73.33-39)

## Вступ

Епілепсія є одним з найпоширеніших неврологічних розладів, що характеризується судомними нападами, які є проявом надмірної гіперсинхронної активності нейронів у мозку. Даний розлад супроводжується збільшеним ризиком смерті, частим розвитком депресії, зниженням когнітивних функцій, працездатності та здатності до навчання. Вказані фактори обумовлюють зниження якості життя хворих та порушення їх трудової та соціальної адаптації.

Основним інструментом у дослідженнях функціонування мозку при епілепсії є вимірювання та аналіз його біоелектричної активності. Електроенцефалографія – це неінвазивний метод моніторингу та дослідження сумарної електрофізіологічної активності і функціонального стану нейронів мозку, який полягає в реєстрації і аналізі різниці потенціалів між точками на поверхні голови – електроенцефалограми (ЕЕГ). Сигнали ЕЕГ по своїй суті є нестационарними часовими рядами, до яких можливо застосувати сукупність математико-статистичних методів аналізу. Одним з найбільш поширених методів аналізу сигналів є побудова математичних моделей та подальший аналіз їх параме-

трів з метою отримання додаткової інформації про сам сигнал.

Авторегресійні моделі широко використовуються для опису сигналів ЕЕГ. Наприклад, в роботі [1] було застосовано часовий та спектральний аналіз мультиваріативних авторегресійних моделей (МВАР) часових рядів для визначення взаєм- та причинно-наслідкового (каузального) зв'язку між магнітоенцефалографічними сигналами або сигналами ЕЕГ. Окрім того, в роботі представлено комплексне визначення різних форм причинності та зв'язків, які можна спостерігати в МВАР моделях, їх розрахунок та інтерпретація. В роботі [2] досліджується закономірність розповсюдження електричної активності між ділянками мозку щурів, специфічних для епілептичних осередків. Показана доцільність використання статистичного моделювання ЕЕГ-сигналів та розрахунку мір для виявлення зв'язків між багатоканальними записами ЕЕГ на основі коефіцієнтів МВАР моделей. В [3] також було показано доцільність розрахунку та використання МВАР-моделей у аналізі причинно-наслідкових зв'язків між ЕЕГ-сигналами. Окрім того, проведено дослідження використання та можливості візуалізації зазначених зв'язків за допомогою використання теорії графів. У моделюванні ЕЕГ на основі

МВАР основним питанням є вибір порядку моделі, яка використовується для опису властивостей безпосередньо сигналу ЕЕГ, а також для подальшого отримання відомостей про функціональні зв'язки в мозку.

Задачею досліджень у цій роботі є застосування МВАР-моделювання до аналізу електричних сигналів мозку пацієнтів, яким поставлений діагноз епілепсії. Зокрема, робота присвячена визначенню та порівнянню порядків МВАР-моделей для ділянок сигналів до, під час та після епілептичного нападу з метою визначення можливих відмінностей у сигналах на цих ділянках та уніфікації підходів до застосування МВАР-моделей в подальших дослідженнях.

## 1 Мультиваріативне авторегресійне моделювання сигналів ЕЕГ

Авторегресійна модель (АР) припускає, що значення дискретного сигналу  $y(n)$  в момент часу  $n$  може бути лінійно спрогнозоване на основі  $p$  минулих значень цього сигналу наступним чином:

$$y(n) = a_1y(n-1) + a_2y(n-2) + \dots + a_p y(n-p) + u(n), \quad (1)$$

де  $p$  є порядком АР моделі,  $a_i$  – коефіцієнти моделі та  $u(n)$  – залишки, що являють собою різницю між початковим часовим рядом та його оцінкою, що виражається АР моделлю. Якщо модель відображає часовий ряд, то залишки – це білий некорельований шум. Це означає, що вся інформація, що міститься в початкових даних, моделюється системою АР.

Розглянемо  $M$  сигналів каналів ЕЕГ – стаціонарних стохастичних процесів  $y_m(n)$ . МВАР моделювання є очевидним розширенням АР моделі. Для моделювання МВАР, рівняння для одного сигналу, наведене вище, можна записати як:

$$y_m(n) = a_{m1,1}y_1(n-1) + a_{m2,1}y_2(n-1) + \dots + a_{mM,1}y_M(n-1) + \dots + a_{mM,p}y_M(n-p) + u_m(n), \quad (2)$$

$m = 1, \dots, M$ , де  $y_m(n)$  – лінійна комбінація його минулих значень та минулих значень інших сигналів  $M-1$ .

Нехай  $Y(n) = [y_1(n), \dots, y_M(n)]'$  – матриця, що містить сигнали, а  $U(n) = [u_1(n), \dots, u_M(n)]'$  містить залишки та коефіцієнти містяться в матриці  $A_k$ :

$$A_k := \begin{bmatrix} a_{11,k} & \dots & a_{1M,k} \\ \vdots & \ddots & \vdots \\ a_{M1,k} & \dots & a_{MM,k} \end{bmatrix}. \quad (3)$$

Квадратна матриця  $A_k$  розміром  $M \times M$  – це матриця коефіцієнтів, що визначає лінійну взаємозалежність між поточними відліками всіх сигналів та відліками всіх сигналів при затримці  $k$ ,  $k = \overline{1:p}$  [1, 3, 4]. Тоді рівняння в матричній формі буде мати такий вигляд:

$$Y(n) = A_1Y(n-1) + A_2Y(n-2) + \dots + A_pY(n-p) + U(n) = \sum_{k=1}^p A(k)Y(n-k) + U(n). \quad (4)$$

МВАР моделі ЕЕГ, в основному, формують розв'язання систем лінійних рівнянь (2) та дозволяють проводити більш глибокі дослідження та аналіз (сегментація ЕЕГ, класифікація ЕЕГ, прогнозування і т.д.) з високою точністю. В цій роботі використовується метод розв'язання систем лінійних рівнянь Юла-Уолкера. Розв'язок даних рівнянь у матричній формі дозволяє за попередньо обраними автокореляціями часових рядів визначити вектор коефіцієнтів оператора регресії у МВАР моделі [5].

## 2 Обирання оптимального порядку МВАР моделі

Від порядку моделі залежить точність представлення часових рядів – каналів ЕЕГ. Оскільки найкраще значення порядку моделі  $p$  заздалегідь є невідомим, то необхідно випробувати декілька значень для вибору оптимального. Для того, щоб це зробити, потрібно ввести той чи інший інформаційний критерій як міру якості статистичної МВАР моделі. Якщо порядок моделі обрано занадто малим, то будуть отримані сильно згладжені спектральні оцінки сигналу, якщо занадто великим – збільшиться роздільна здатність, але з'являться помилкові спектральні піки [5].

Після опрацювання існуючих методів оцінки якості МВАР статистичних моделей, таких як інформаційний критерій Акаїке, критерій Ханнана-Квіна та ін., було обрано інформаційний критерій Шварца-Байеса (Schwarz-Bayes Criterion, *SBC*), оскільки він тісно пов'язаний як з критерієм Акаїке, так і з критерієм Ханнана-Квіна [2, 4, 6, 7]. Основною різницею між ними є те, що при використанні *SBC* модель більш суворо штрафується за збільшення кількості її параметрів, тому існує можливість обрати найбільш оптимальний порядок для моделі за заданих реальних даних [8, 9]. При обиранні з декількох моделей найприйнятнішою є модель із найнижчим значенням *SBC*.

Згідно з визначенням, наведеними в [2, 4, 6] критерій *SBC* визначається як:

$$SBC(p) = \ln |\det(S(p))| + \frac{\ln(N)}{N} pM^2, \quad (5)$$

де  $\ln |det(S(p))|$  – це логарифм визначника дисперсійно-коваріаційної матриці залишків  $S$  (помилка прогнозування, що обчислюється на основі зазначеної матриці залишків  $U$ ) для МВАР моделі порядку  $p$ ,  $N$  – кількість відліків сигналу та  $M$  – кількість досліджуваних сигналів. Визначник матриці  $S$  є так званою узагальненою дисперсією, яка кількісно характеризує ступінь випадкового розсіювання значень векторної випадкової величини навколо свого середнього. Очевидно, що при збільшенні кількості параметрів моделі перша складова в правій частині зменшується, а друга збільшується.

Якість побудованої моделі в цій роботі також запропоновано оцінювати за допомогою коефіцієнта кореляції ( $CC$ ) Пірсона між початковим часовим рядом та його оцінкою, що виражається за допомогою коефіцієнтів МВАР моделі. Окрім того, також було розглянуто можливості використання інших статистичних методів оцінки якості моделі, таких як критерій Льюнга-Бокса, критерій Дарбіна-Уотсона, тест Бройша-Годфрі та інші [2, 4]. Використання даних тестів планується дослідити в подальшій роботі над досліджуваною проблемою.

Окрім МВАР моделі, яка представлена матрицею  $A$ , що наведена в рівнянні (3) і визначає лінійну взаємозалежність між сигналами, також обчислюються ще три матриці на основі матриці коефіцієнтів  $A$ . Це матриця залишків  $U$  розміром  $M \times (N - p)$ , коваріаційна матриця залишків  $S$  розміром  $M \times M$  та матриця  $Y$  змодельованих сигналів, розміром  $M \times (N - p)$ . Якщо модель добре підходить до реальних даних ЕЕГ, то отримані сигнали повинні практично їм відповідати.

### 3 Постановка експерименту та аналіз результатів

#### 3.1 Опис використаних даних

Для вирішення задачі дослідження розповсюдження та взаємозв'язку біоелектричної активності різних ділянок мозку було використано записи ЕЕГ-сигналів з проявами епілептичної активності [10–12]. Сигнали ЕЕГ були зареєстровані у медичному закладі під наглядом лікаря протягом нічного моніторингу. Було використано два набори даних з генералізованими та фокальними епілептичними нападами. Перша група пацієнтів з фокальною формою нападів складається з 26 осіб, у яких разом було зареєстровано більше 100 епілептичних нападів. Друга група з генералізованими епілептичними нападами складається з 11 осіб, у яких разом зареєстровано близько 50 епілептичних нападів. ЕЕГ була зареєстрована з частотою 250 Гц за схемою 10/20. Епізоди початку і закінчення нападу білі розмічені лікарем на основі візуального спостереження та подальшого аналізу відеозаписів. На рис. 1

наведені ділянки ЕЕГ-сигналів пацієнта з характерно вираженим епілептичним нападом.

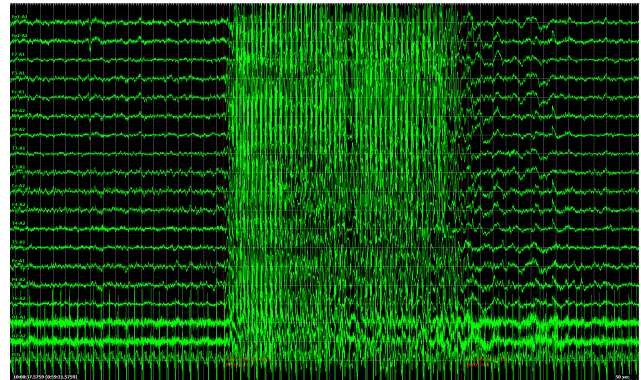


Рис. 1. Приклад запису ЕЕГ-сигналу, що містить епілептичний напад

#### 3.2 Побудова МВАР моделей ЕЕГ

Для дослідження причинно-наслідкового зв'язку між ЕЕГ-сигналами під час різних станів пацієнтів з епілепсією, зокрема до, під час та після епілептичного нападу, було вирішено провести їх статистичне моделювання в середовищі MATLAB на основі алгоритму, розробленого в роботі [1]. Були створені МВАР моделі для зазначених наборів даних та для усіх епілептичних нападів, а також для ділянок сигналів до та після нападу. Оскільки на даний час немає рекомендацій щодо оптимальної для дослідження тривалості сигналів до та після нападу, в даній роботі їх було обрано в два рази більшими, ніж тривалість відповідного нападу. Для обчислення моделей використовувались значення порядків від 1 до 22, тобто для кожного досліджуваного сигналу будуються та порівнюються 22 різні моделі.

#### 3.3 Дослідження МВАР моделей для різних ділянок ЕЕГ

Після проведення моделювання, отримані МВАР моделі для кожного значення порядку були оцінені за допомогою критерію  $SBC$ . Окрім того, були обчислені значення  $CC$  між реальними та змодельованими сигналами. На рис. 2 наведено приклад визначення оптимальних порядків обчислених МВАР моделей в залежності від мінімальних значень критерію  $SBC$  для ділянки з нападом у типового пацієнта.

Окрім того, на рис. 3 наведено графік залежності  $CC$  між реальним та змодельованим сигналом ЕЕГ від порядку моделі МВАР ділянки сигналів, що містять напади. Як видно, порядок, обраний за допомогою критерію, також відповідає коефіцієнту кореляції не менше ніж 0,995. Це додатково свідчить про те, що порядок моделі МВАР достатній для отримання точного моделювання.  $CC$  не являється

таким інформативним, як критерій  $SBC$ , оскільки, як було зазначено у формулі (5),  $SBC$  оцінює коваріаційну матрицю залишків, які являють собою похибку моделювання; але  $CC$  додатково свідчить про те, що модель відображає ЕЕГ сигнали.

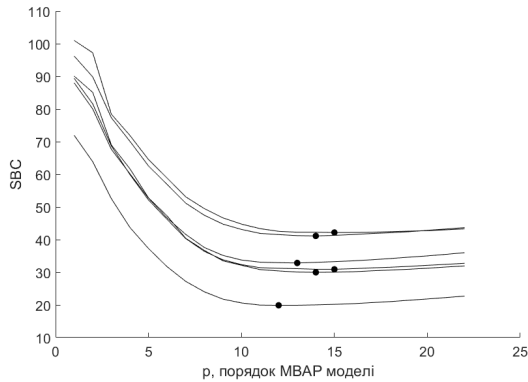


Рис. 2. Залежності критерію  $SBC$  від порядку МВАР моделей ЕЕГ під час епілептичних нападів для типового пацієнта

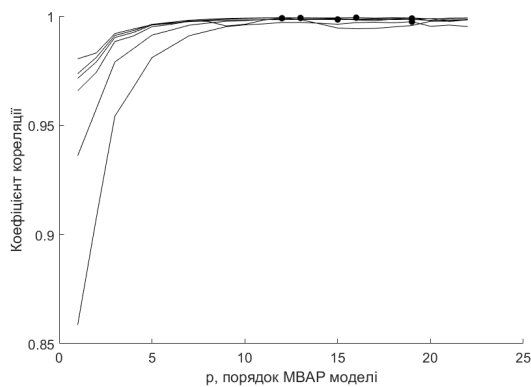


Рис. 3. Залежності коефіцієнту кореляції Пірсона між реальними та змодельованими ділянками ЕЕГ сигналів під час нападу від порядку моделі для типового пацієнта

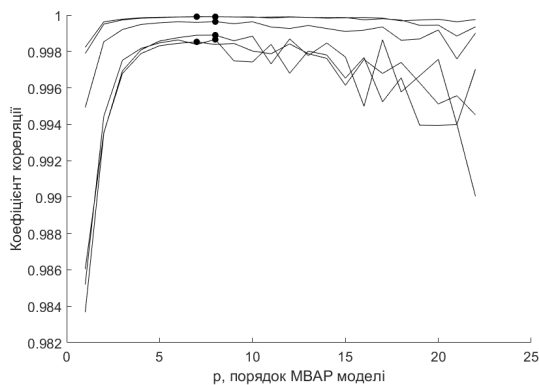


Рис. 4. Залежності коефіцієнту кореляції Пірсона між реальними та змодельованими ділянками ЕЕГ сигналів до нападу від порядку моделі для типового пацієнта

Слід також зазначити, що, згідно з отриманими результатами,  $CC$  для вищих порядків МВАР моделей ділянок сигналів під час епілептичного нападу має більш детерміновану природу, ніж  $CC$  аналогічних порядків для ділянок до та після нападу, які є більш стохастичними (рис. 4). Ми пов'язуємо це з тим, що сигнали ЕЕГ під час нападу містять характерні ознаки, такі як високоамплітудні низькочастотні коливання та комплекси спайк-хвиля [10]. Тому, в силу наявності цих періодичних патернів, моделювання сигналів з епілептичними нападами виявляється більш точним з використанням значень порядків у вказаному діапазоні.

На рис. 5 для ілюстрації наведені графіки сигналу ЕЕГ та його моделі. Можемо візуально переконатись, що отримано модель, яка достатньо точно відповідає початковому сигналу ЕЕГ.

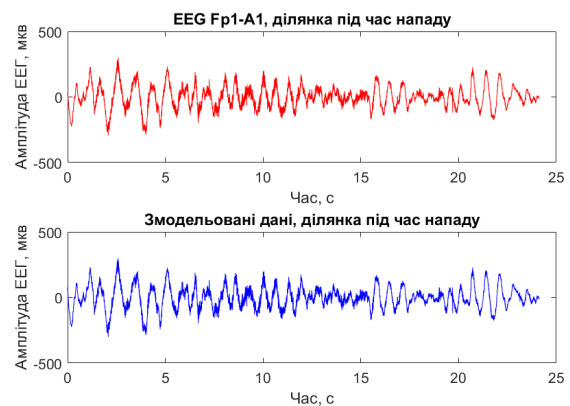


Рис. 5. Вимірний сигнал ЕЕГ (верхній графік) та його модель на основі коефіцієнтів МВАР (нижній графік)

Точками на рис. 2-4 позначені значення оптимальних порядків. Можна констатувати, що в результаті роботи отримано МВАР моделі сигналів ЕЕГ та проведено аналіз якості розрахованих моделей залежно від їх порядку.

### 3.4 Статистичний аналіз результатів

В результаті проведеної роботи були обчислені оптимальні значення порядків МВАР моделей ділянок ЕЕГ сигналів з фокальними та генералізованими нападами. МВАР моделі були обчислені для ділянок сигналів до, під час та після нападу та оцінені за допомогою  $CC$  та критерію  $SBC$ . Для кожного з зазначених наборів даних були сформовані три вибірки, кожна з яких містить значення оптимальних порядків. Вибірки є зв'язними, тобто відповідні порядки відповідають одному і тому ж пацієнту, проте під час різних станів активності його мозку.

Після застосування критерію узгодженості Колмогорова-Смирнова до кожної вибірки з двох наборів даних, було відхилено нульову гіпотезу про те, що вибірки підпорядковуються нормальному

закону розподілу на рівні значущості  $p = 0.05$ . Окрім того, після застосування даного критерію для попарно зв'язаних вибірок, було визначено, що вибірки значень порядків для фокальних епілептичних нападів *до* та *під час* нападу і *до* та *після* нападу мають однаковий розподіл. Вибірки порядків *під час* та *після* нападу мають різний розподіл. Для генералізованих нападів даний тест свідчить про однаковий розподіл усіх вказаних вибірок.

Після відхилення гіпотези про нормальний розподіл даних було вирішено застосувати критерій Вілкоксона для зв'язаних вибірок (Wilcoxon signed-rank test) – непараметричний статистичний критерій, що застосовується для оцінки відмінностей між двома залежними вибірками, взятими із закону розподілу, відмінного від нормального.

За результатами даного тесту при фокальній формі нападу, коли патологічна активність поширюється лише по одній ділянці мозку, було підтверджено нульову гіпотезу про статистичну відмінність між вибірками порядків *до* та *після* нападу на рівні значущості  $p = 0.0769$ . Тобто можна досліджувати природу залежності порядків МВАР моделей для цих ділянок сигналів окремо. Щодо вибірок порядків МВАР моделей *до* та *під час* нападу і *під час* та *після* нападу, то статистичної відмінності не було виявлено на рівнях значущості  $p = 0.004$  та  $p = 1.84 \cdot 10^{-15}$  відповідно.

При генералізованій формі, коли електрична активність під час нападу поширюється по всьому мозку, критерій Вілкоксона показав протилежні результати. Не було відхилено нульову гіпотезу про статистичну відмінність між вибірками порядків МВАР-моделей сигналів *до* та *під час* нападу і *під час* та *після* нападу на рівнях значущості  $p = 0.542$  та  $p = 0.1263$  відповідно. Між вибірками порядків сигналів *до* та *після* нападу не було виявлено статистичної відмінності на рівні значущості  $p = 0.0041$ .

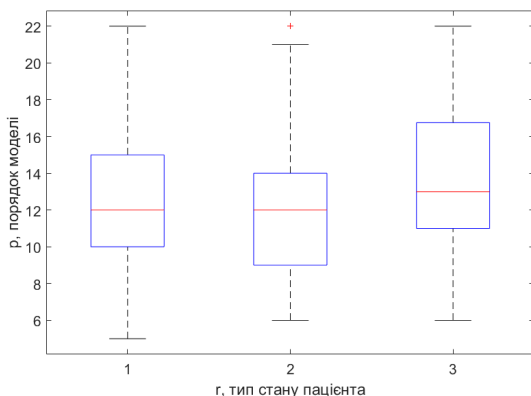


Рис. 6. Коробчаті діаграми для порівняння медіан порядків МВАР моделей сигналів при фокальній формі нападів, 1–ділянка до нападу, 2–під час нападу, 3–після нападу

На рис. 6 наведені коробчаті діаграми для порівняння медіан обчислених порядків моделей дослід-

жуваних сигналів при фокальній формі нападів. Можемо стверджувати, що в основному порядки обираються рівними 12 для сигналів, що передують нападам та містять напади і рівними 13 для сигналів після нападів.

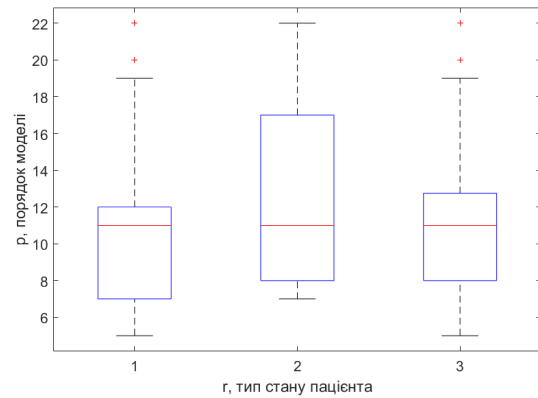


Рис. 7. Коробчаті діаграми для порівняння медіан порядків МВАР моделей сигналів при генералізованій формі нападів, 1–ділянка до нападу, 2–під час нападу, 3–після нападу

При генералізованій формі епілепсії (рис. 7) можемо спостерігати схожу картину. Медіани порядків для усіх типів досліджуваних сигналів дорівнюють 11.

## Висновки

Отже, на основі отриманих результатів можемо зробити висновок щодо вибору значення оптимального порядку при МВАР-моделюванні ЕЕГ сигналів, що містять епілептичні напади. Рекомендовано величину порядку обирати в діапазоні значень від 11 до 13. Оскільки частота дискретизації використаних сигналів складає 250 Гц, то вказаний діапазон значень порядків свідчить про те, що МВАР-моделювання сигналів відбувається з урахуванням інформації, яку містять усі інші сигнали при затримці до 44-52 мс. Тобто на основі коефіцієнтів отриманих МВАР-моделей наших сигналів, теоретично можемо виділяти функціональні характеристики електричної активності нейронів мозку при епілепсії, що виникають синхронно у різних ділянках мозку та розповсюджуються, в середньому, протягом 50 мс. Окреме моделювання ділянок сигналів *до* та *під час* нападу може дати змогу краще зрозуміти природу виникнення самого нападу та важливі функціональні характеристики електричної активності мозку, що йому передують.

Отримані МВАР-моделі сигналів можна використовувати в якості параметрів для алгоритмів машинного навчання та нейронних мереж для створення систем прогнозування епілептичних нападів. Подальша робота над досліджуваною проблемою буде стосуватися розрахунку оптимального набору

параметрів для характеристики зазначених функціональних зв'язків. Набір параметрів може включати в себе *когерентність*, що є аналогом коефіцієнта взаємної кореляції між сигналами в частотній області; *направлену когерентність*, що слугує інструментом для вивчення функціональних взаємодій між структурами мозку; *часткову когерентність*, яка дозволяє виявити лінійні впливи між сигналами; *часткову спрямовану когерентність*, яка описує відносини (напрямок потоку інформації) між структурами мозку; *функцію спрямованої передачі*, яка описує потік інформації між сигналами в залежності від частоти, та деякі інші. Всі перераховані параметри обчислюються з використанням отриманих матриць коефіцієнтів МВАР-моделей, обрання та розрахунок яких було обґрунтовано в даній роботі для сигналів, що містять епілептичні напади.

Дослідження даної проблеми є актуальним і важливим для розуміння природи розповсюдження електричної активності між нейронами мозку *до, під час* та *після* нападу, та для прогнозування самих нападів, засновуючись на ділянках сигналів, що їм передують.

## Перелік посилань

1. Faes L. Testing Frequency Domain Analysis of Causal Interactions in Physiological Time Series / L. Faes, G. Nollo. – IEEE. – 2010. – No 57(8). – pp. 1897-1906.
2. Van Mierlo P. Changes in connectivity patterns in the kainate model of epilepsy / P. van Mierlo, S. Asseondi, P. Boon, I. Lemahieu, eds. – Berlin: Springer. – 2009. – pp. 35-54.
3. Seth A. Causal connectivity of evolved neural networks during behavior / A. Seth. – Network: Computation in Neural Systems. – 2005. – № 16(1). – pp. 360-363
4. Lütkepohl H. New introduction to multiple time series analysis / A. Massaro. – Berlin: Springer. – 2005. – 764 p.
5. Жаринов И.О. К вопросу о выборе порядка авторегрессионных моделей сигналов электроэнцефалограмм человека (в медицинском приборостроении) / И.О. Жаринов // Научно-технический вестник информационных технологий, механики и оптики. – 2006. – № 33. – 12 p.
6. Hurvich C. Regression and Time Series Model Selection in Small Samples / C. Hurvich, CL. Tsai // Biometrika. – 1989. – Vol 76, No 2. – pp. 297.
7. Porcaro C. Choice of multivariate autoregressive model order affecting real network functional connectivity estimate / C. Porcaro, F. Zappasodi, PM. Rossini, F. Tecchio // Biometrika. – 2009. – Vol. 120, No 2. – pp. 436-448.
8. Chatfield C. The Analysis of Time Series: An Introduction / C. Chatfield. – CRC Press, 1980. – 11 p.
9. Vrieze S. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) / S. Vrieze // Psychological Methods. – 2012. – Vol. 17, No 2. – pp. 228-243.

10. Панічев О. Ю. Порівняння результатів прогнозування епілептичних нападів при використанні різних схем відведення ЕЕГ / О.Ю. Панічев, А.О. Попов, В.І. Харитонов // Вісник НТУУ «КПІ». Серія Радіотехніка. Радіоапаратобудування. – 2017. – № 68. – с. 54-58.
11. Popov A. Heart beat-to-beat intervals classification for epileptic seizure prediction / A. Popov, O. Panichev, Y. Karplyuk, Y. Smirnov, S. Zaunseder and V. Kharytonov // SPSympo. – 2017
12. Smirnov Y. Epileptic seizure prediction based on singular value decomposition of heart rate variability features / A.Y. Smirnov, A. Popov, O. Panichev, Y. Karplyuk and V. Kharytonov // SPSympo. – 2017.

## References

- [1] Faes L., Porta A. and Nollo G. (2010) Testing Frequency-Domain Causality in Multivariate Time Series. *IEEE Transactions on Biomedical Engineering*, Vol. 57, Iss. 8, pp. 1897-1906. DOI: 10.1109/tbme.2010.2042715
- [2] Van Mierlo P., Asseondi S., Staelens S., Boon P. and Lemahieu I. (2009) Changes in connectivity patterns in the kainate model of epilepsy. *IFMBE Proceedings*, pp. 360-363. DOI: 10.1007/978-3-540-89208-3\_85
- [3] Seth A.K. (2005) Causal connectivity of evolved neural networks during behavior. *Network: Computation in Neural Systems*, Vol. 16, Iss. 1, pp. 35-54. DOI: 10.1080/09548980500238756
- [4] Lütkepohl H. (2005) Introduction. *New Introduction to Multiple Time Series Analysis*, Springer, pp. 1-7. DOI: 10.1007/978-3-540-27752-1\_1
- [5] Zharinov I.O. (2006) К вопросу о выборе порядка авторегрессионных моделей сигналов электроэнцефалограмм человека (в медицинском приборостроении) [On the choice of the order of autoregressive models of signals of human electroencephalograms]. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, No 33, p. 121-132.
- [6] Hurvich C.M. and Tsai C. (1989) Regression and Time Series Model Selection in Small Samples. *Biometrika*, Vol. 76, Iss. 2, pp. 297. DOI: 10.2307/2336663
- [7] Porcaro C., Zappasodi F., Rossini P.M. and Tecchio F. (2009) Choice of multivariate autoregressive model order affecting real network functional connectivity estimate. *Clinical Neurophysiology*, Vol. 120, Iss. 2, pp. 436-448. DOI: 10.1016/j.clinph.2008.11.011
- [8] Chatfield C. (1980) Introduction. *The Analysis of Time Series: An Introduction*, CRC Press, pp. 1-11. DOI: 10.1007/978-1-4899-2923-5\_1
- [9] Vrieze S.I. (2012) Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, Vol. 17, Iss. 2, pp. 228-243. DOI: 10.1037/a0027127
- [10] Panichev O., Popov A. and Kharytonov V. (2017) Comparison of epileptic seizure prediction performance for different EEG derivation schemes. *Visn. NTUU KPI, Ser. Radiotekh. radioaпаратobuduv.*, Vol. 68, pp. 54-58. DOI: 10.20535/radap.2017.68.54-58
- [11] Popov A., Panichev O., Karplyuk Y., Smirnov Y., Zaunseder S. and Kharytonov V. (2017) Heart beat-to-beat intervals classification for epileptic seizure prediction. *2017 Signal Processing Symposium (SPSymo)*. DOI: 10.1109/sps.2017.8053647

- [12] Smirnov Y., Popov A., Panichev O., Karplyuk Y. and Kharytonov V. (2017) Epileptic seizure prediction based on singular value decomposition of heart rate variability features. *2017 Signal Processing Symposium (SPSymposium)*. DOI: 10.1109/sps.2017.8053648

## Выбор оптимального порядка мультивариативных авторегрессионных моделей электроэнцефалограмм при эпилепсии

*Котючий И. В., Попов А. О., Харитонов В. И.*

В работе предложен подход к выбору оптимальных порядков мультивариативных авторегрессионных (МВАР) моделей сигналов электрической активности мозга пациентов, которым поставлен диагноз эпилепсии. Проведено МВАР-моделирование и статистический анализ оптимальных порядков моделей участков сигналов до, во время и после эпилептического приступа. Для оценки качества моделей использован информационный критерий Шварца-Байеса с учетом ковариационной матрицы остатков смоделированных данных. Дополнительно качество оценено с помощью коэффициента корреляции Пирсона между реальными и смоделированными данными. В экспериментальной части были использованы записи электроэнцефалографических (ЭЭГ) сигналов с двух наборов данных для генерализованных и фокальных эпилептических приступов. По результатам проведенной работы рекомендуется выбирать значение порядков МВАР моделей ЭЭГ-сигналов для указанных участков в диапазоне 11-13. Указаны пути дальнейших исследований электрической активности мозга и функциональных связей его отделов во время эпилептической активности.

*Ключевые слова:* авторегрессионная модель; электроэнцефалография; эпилепсия; порядок статистической модели; эпилептические приступы

## Selection of the optimal order for multivariate autoregressive model of electroencephalograms for patients with epilepsy

*Kotiuchyi, I. V., Popov, A. O., Kharytonov, V. I.*

**Introduction.** Brain electrical activity signals (or EEG) by their very nature are non-stationary time series. This basically allows applying a set of mathematical-statistical analysis methods to them. One of the most common methods for signal analyzing is the construction of autoregressive mathematical models and analysis of their

parameters in order to obtain additional information about the signal itself or causality between signals. In multivariate autoregressive (MVAR) modeling of EEG, the main issue is the optimal choice of model order. In this work, the approach for selecting the optimal order of MVAR models of brain electrical activity signals of subjects diagnosed with epilepsy is proposed.

**MVAR modeling.** The autoregressive model assumes that the current sample of the discrete signal can be linearly predicted as a weighted sum of its previous samples. MVAR model extends this assumption to multiple time series so that the vector of current samples of all signals is modeled as a linear sum of their previous samples. MVAR models of EEG signals essentially are formed by solving systems of linear equations. The Yule-Walker method of linear equations systems solving is used in this paper. The accuracy of EEG modeling depends on the order of model. Each model order influenced by the amount of delay between current samples and last previous samples used to generate the model. To assess the order and quality of models the Schwarz-Bayes information criterion (SBC) is used in this work taking into account the covariance matrix of the residuals. Additionally, the quality is assessed by Pearson's correlation coefficient between the real and simulated data. In this paper, the MVAR modeling and statistical analysis of the models' optimal orders of the input signal periods before, during and after an epileptic seizure is carried out.

**Experimental results.** Two sets of EEG data with generalized and focal epileptic seizures are used. The first group of patients with focal seizures consists of 26 people and more than 100 epileptic seizures. The second group with generalized seizures consists of 11 people and about 50 epileptic seizures. For EEG signals modeling, values of orders in a range from 1 to 22 are used. Consequently, for each investigated period of signal (before, during and after a seizure), 22 different MVAR models are constructed and compared. After modeling, the obtained models for each order value are evaluated using the SBC criterion.

**Conclusions.** According to the results, it is recommended to choose the order of MVAR models of EEG signals in the predefined range of orders from 11 to 13. Since the sampling rate of the signals used in these experiments is 250 Hz, the specified range of order values indicates that MVAR-modelling of one signal includes information that contains all other signals with a delay of 44-52 ms. Therefore, theoretically, it is possible to allocate functional characteristics of brain electrical activity for patients with epilepsy that occur synchronously in different parts of the brain and spread at an average of 50 ms. Moreover, the ways of further research of electrical brain activity and functional connections of brain regions during epileptic activity are indicated.

*Key words:* autoregressive model; electroencephalography; epilepsy; the order of statistical model; epileptic seizures